ASSESSMENT AND PREDICTION OF BREAST CANCER OUTCOME IN SUBGROUPS OF PATIENTS

MIAO HUI

(BSc Hons, National University of Singapore)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

SAW SWEE HOCK SCHOOL OF PUBLIC HEALTH

NATIONAL UNIVERSITY OF SINGAPORE

2015

Declaration

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Miao Hui

January, 2015

Acknowledgements

This PhD thesis would not have been possible without the help and support from my supervisors, mentors, colleagues, friends and family.

I am more than lucky to have two amazing supervisors A/P Mikael Hartman and A/P Helena Verkooijen. Their enthusiasm and determination towards research and education always inspire me to step outside my comfort zone and dream big for my future. I was given numerous opportunities to explore my own interest and gain greater exposure such as studying and working overseas at UMC Utrecht and collaboration with the largest international breast cancer consortium BCAC. All these experiences have made my PhD journey full of excitement and fulfillment. I am very grateful of all their guidance and advice on both professional and personal level. They are indeed my role models I can look up to. Tack så mycket, Prof Hartman and dank je wel, Prof Lenny, for paving the pathway and shaping my future to become a real epidemiologist of tomorrow.

I also would like to express my sincere gratitude to Prof Chia Kee Seng, Dean of Saw Swee Hock School of Public Health. I still remembered my first meeting with Prof Chia when I was still an undergraduate statistics student who had no clue what epidemiology actually meant. Impressed by his passion and his vision of epidemiology research in Singapore, I decided to apply for PhD programme. During my studies, he always took time out of his busy schedule to interact with fellow PhD students and I enjoyed all his thought-provoking speeches.

I would like to give special thanks to Prof Yip Cheng Har, Dr Nirmala Bhoo Pathy, Dr Nur Aishah Taib, Dr Tan Ern Yu and all other collaborators from NUH, UMMC

and TTSH, for their contribution to the Singapore Malaysia Working Group. Their drive to improve breast cancer care and research in Asia is always my motivation. I must thank Dr Nirmala for always replying my emails promptly and spending hours with me on skype. I really appreciate her insightful comments and suggestions.

It is great honor to receive professional advice from my Thesis Advisory Committee chairperson A/P Tai Bee Choo and my mentor for my NMRC fellowship A/P Teo Yik Yin. Their dedication to research has encouraged me to pursue my passion in statistics.

To our breast cancer research team, Jenny, Siew Li, Siok Hoon, Ting Ting, Amanda, Yi Zhen, Peh Joo, Charmaine, Dr Li Jingmei, and all other colleagues from Saw Swee Hock School of Public Health who have helped me in one way or another, I could not ask for a more fun-loving, helpful and enthusiastic group of people to work with. I want to thank them for making my work life at SPH a truly unforgettable one.

I am very grateful to NMRC for sponsoring my PhD training and overseas travel.

Last but not least, I would like to thank my husband and my parents who have always been my pillars of support and believed in me for who I really am.

Table of Contents

Summary

Global incidence rate of breast cancer is on the rise and according to WHO GLOBOCAN report, today about 1.7 million women are diagnosed with the disease annually worldwide [1, 2]. In Singapore, breast cancer is the most common female cancer, accounted for approximately 29% of all cancer cases and attributed to 17% of cancer deaths among women in 2003-2007 [3].

Breast cancer is a heterogeneous disease in terms of histopathological characteristics, response to treatment and clinical outcome, which makes prognostication rather challenging. Survival has been improved in many developed countries over the last decades, partially owing to mammographic screening and improvement in diagnostic tools and treatment modalities. The survival rate in many developing countries in Asia, however, remains relatively poor and prognostic tools developed in Western countries were shown to be inadequate or inappropriate for Asian population. In this thesis, we propose four studies for different subgroups of patients to assess their outcome, evaluate prognostic factors and validate prognostic models.

Many risk and prognostic factors have been identified for female breast cancer but how they affect disease risk and survival of male patients is uncertain. In the first study of the thesis, we investigated breast cancer, diagnosed between 1970 and 2007 from six population-based cancer registries. We found that the incidence of male breast cancer remained at a stable and low rate. However male patients presented with more advanced stages than women. After adjustment for age, stage and treatment, men had better relative survival than women.

The introduction of nationwide screening programme for breast cancer has dramatically increased the incidence of breast carcinoma *in situ*, a non-invasive malignant lesion. The second study in this thesis assessed prognosis of 8111 women with *in situ* breast cancer registered between 1980 and 2004 in the population-based Swedish Multi-Generation Register. Women with carcinoma *in situ* were four times more likely to develop invasive breast cancer compared to women in the general population and the excess risk was more pronounced in younger women and those with family history.

In the third study in this thesis, we validated CancerMath, a web-based prognostic tool, in patients with stage I to stage III breast cancer registered between 1990 and 2011 in the Singapore Malaysia Hospital Based Breast Cancer Registry. Discrimination and calibration of CancerMath was modest and prediction was more accurate for patients with favorable tumor characteristics. In the fourth study, a systematic review was conducted to identify existing prognostic tools for *de novo* metastatic breast cancer. We validated nine out of 16 models in 642 Asian women with *de novo* metastatic breast cancer diagnosed from 2000 to 2010 in the Singapore Malaysia Hospital Based Breast Cancer Registry. The discriminatory performance of these models was modest. The third and fourth studies suggest that development of Asian-specific prediction tools is needed to improve prognostication and to guide decision making.

List of Tables

List of Figures

List of Abbreviations

AC	Doxorubicin and cyclophosphamide
AI	Aromatase inhibitors
AJCC	American Joint Committee on Cancer
ALKP	Alkaline phosphatase
ALND	Axillary lymph node dissection
ASRS	Age-standardized relative survival
AUC	Area under the curve
CEP17	Chromosome enumeration probe 17
CI	Confidence interval
CK	Cytokeratin
CMF	Cyclophosphamide methotrexate and fluorouracil
CTC	Circulating tumor cell
C-statistic	Concordance statistic
DCIS	Ductal carcinoma *in situ*
EAR	Excess additive risk
ECOG	Eastern Cooperative Oncology Group
EGFR	Epidermal growth factor receptor
ER	Estrogen receptor
FDA	Food and Drug Administration
FEC	Fluorouracil, epirubicin and cyclophosphamide
FISH	Fluorescence in situ hybridization
GWAS	Genome-wide association studies
HER2	Human epidermal growth factor 2
HRT	Hormone replacement therapy
IARC	International Agency for Research on Cancer
IDC	Invasive ductal carcinoma
IHC	Immunohistochemistry
ILC	Invasive lobular carcinoma
KPS	Karnofsky performance status
LCIS	Lobular carcinoma *in situ*
LDH	Lactate dehydrogenase
MGR	Multi-Generation Register
MR	Mortality ratio

NPI	Nottingham Prognostic Index
NRDO	National Registry of Disease Office
NUH	National University Hospital
pCR	Pathological complete response
PR	Progesterone receptor
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-analyses
PROM	Patient-reported outcome measure
RCT	Randomized controlled trial
RER	Relative excess risk
ROC	Receiver operating characteristic
RSR	Relative survival ratio
SEER	Surveillance, Epidemiology, and End Results
SIR	Standardized incidence ratio
SLN	Sentinel lymph node
SMR	Standardized mortality ratio
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TAC	Docetaxel, doxorubicin and cyclophosphamide
TTSH	Tan Tock Seng Hospital
UMMC	University Malaya Medical Centre
WHO	World Health Organization

List of Publications

This thesis is based on the following publications/manuscripts:

1. Incidence and outcome of male breast cancer: An international population-based study
   Miao H, Verkooijen HM, Chia KS, Bouchardy C, Pukkala E, Larønningen S, Mellemkjær L, Czene K, Hartman M.
   Journal of Clinical Oncology 2011 Nov 20; 29 (33):4381-6

2. Predicting survival of *de novo* metastatic breast cancer in Asian women: systematic review and validation study.
   Miao H, Hartman M, Bhoo-Pathy N, Lee SC, Taib NA, Tan EY, Chan P, Moons KG, Wong HS, Goh J, Rahim SM, Yip CH, Verkooijen HM.
   PLoS One. 2014 Apr 2;9(4):e93755. doi: 10.1371/journal.pone.0093755.

3. The impact of *in situ* breast cancer and family history on risk of subsequent breast cancer events and mortality -a population-based study from Sweden
   Sackey H, Miao H, Czene K, Verkooijen HM, Edgren G, Frisell J, Hartman M
   Submitted to Breast Cancer Research and Treatment

4. Validation of the CancerMath prognostic tool for breast cancer in Southeast Asia
   Hui Miao, Mikael Hartman, Helena M Verkooijen, Nur Aishah Taib, Hoong-Seam Wong, Shridevi Subramaniam, Cheng-Har Yip, Ern-Yu Tan, Soo-Chin Lee, Nirmala Bhoo-Pathy
   Submitted to British Journal of Cancer

Other relevant publications:

1. Fine-Scale Mapping of the 5q11.2 Breast Cancer Locus Reveals at Least Three Independent Risk Variants Regulating MAP3K1.
   Glubb DM,…, Miao H,...for the Breast Cancer Association Consortium
   Am J Hum Genet. 2014 Dec 17. pii: S0002-9297(14)00475-3.

2. Prediction of positive resection margins in patients with non-palpable breast cancer.
   Barentsz MW, Postma EL, van Dalen T, van den Bosch MA, Miao H, Gobardhan PD, van den Hout LE, Pijnappel RM, Witkamp AJ, van Diest PJ, van Hillegersberg R, Verkooijen HM.
   Eur J Surg Oncol. 2014 Sep 2. pii: S0748-7983(14)01056-7

3. Evidence that breast cancer risk at the 2q35 locus is mediated through IGFBP5 regulation.
   Ghoussaini M…, Miao H,...for the Breast Cancer Association Consortium
   Nat Commun. 2014 Sep 23;4:4999. doi: 10.1038

4. Identification and characterisation of novel associations in the CASP8/ALS2CR12 region on chromosome 2 with breast cancer risk.
   Lin WY…, Miao H,...for the Breast Cancer Association Consortium
   Hum Mol Genet. 2014 Aug 28. pii: ddu431.

5. Genome-wide association analysis in East Asians identifies breast cancer susceptibility loci at 1q32.1, 5q14.3 and 15q26.1.
   Cai Q, Miao H,...for the Breast Cancer Association Consortium
   Nat Genet. 2014 Aug;46(8):886-90. doi: 10.1038/ng.3041.

6. Common non-synonymous SNPs associated with breast cancer susceptibility: findings from the Breast Cancer Association Consortium.
   Milne RL, …, Hui M,...for the Breast Cancer Association Consortium
   Human Molecular Genetics. 2014 Jun 18. pii: ddu311.

7. Pregnancy during breast cancer: does a mother's parity status modify an offspring's mortality risk?
   Simonella L, Verkooijen HM, Edgren G, Liu J, Hui M, Salim A, Czene K, Hartman M.
   Breast Cancer Res Treat. 2014 Jun 17

8. Genetic variation at CYP3A is associated with age at menarche and breast cancer risk: a case-control study.
   Johnson N,…,Miao H,...for the Breast Cancer Association Consortium
   Breast Cancer Research. 2014 May 26;16(3):R51.

9. Recurrent mutation testing of BRCA1 and BRCA2 in Asian breast cancer patients identify carriers in those with presumed low risk by family history.
   Kang PC, Phuah SY, Sivanandan K, Kang IN, Thirthagiri E, Liu JJ, Hassan N, Yoon SY, Thong MK, Hui M, Hartman M, Yip CH, Mohd Taib NA, Teo SH
   Breast Cancer Res Treat. 2014 Mar 1.

10. FGF receptor genes and breast cancer susceptibility: results from the Breast Cancer Association Consortium.
    Agarwal D,…, Hui M,...for the Breast Cancer Association Consortium
    Br J Cancer. 2014 Feb 18;110(4):1088-100. doi: 10.1038/bjc.2013.769.

11. Fine-Scale Mapping of the FGFR2 Breast Cancer Risk Locus: Putative Functional Variants Differentially Bind FOXA1 and E2F1.
    Meyer KB,…,Hui M,...for the Breast Cancer Association Consortium
    The American Journal of Human Genetics 2013 Dec 5;93(6):1046-60.

12. Genome-wide association studies identify four ER-negative specific breast cancer risk loci.
    Garcia-Closas M,…,Miao H,…for the Breast Cancer Association Consortium
    Nature Genetics 2013 Apr;45(4):392-8, 398e1-2.

13. Multiple independent TERT variants associated with telomere length and breast cancer risk
    Bojesen SE,…, Miao H, …for the Breast Cancer Association Consortium
    Nature Genetics 2013 Apr;45(4):371-84, 384e1-2

14. Large-scale genotyping identifies 41 new loci associated with breast cancer risk
    Michailidou K,…, Miao H,…for the Breast Cancer Association Consortium
    Nature Genetics 2013 Apr;45(4):353-61, 361e1-2

15. Functional Variants at the 11q13 Risk Locus for Breast Cancer Regulate Cyclin D1 Expression through Long-Range Enhancers
    French J,…, Miao H,.. for the Breast Cancer Association Consortium
    The American Journal of Human Genetics, 2013 Apr 4;92(4):489-503

16. Common genetic determinants of breast-cancer risk in East Asian women: a collaborative study of 23 637 breast cancer cases and 25 579 controls
    Zheng W,…,Miao H,...for the Breast Cancer Association Consortium
    Human Molecular Genetics. 2013 Mar 27

17. Birth rates among female cancer survivors: a population-based cohort study
    Hartman M, Jenny L, Kamila C, Miao H, Chia KS, Agus S, HM Verkooijen
    Cancer 2013 May 15;119(10):1892-9

18. Ability to predict breast cancer in Asian women using a polygenic susceptibility model.
    Hartman M, Suo C, Lim WY, Miao H, Teo YY, Chia KS.
    Breast Cancer Res Treat. 2011 Jun;127(3):805-12.

Chapter 1 Background

This chapter of the thesis summarizes what is known about breast cancer in terms of disease burden, classification, treatment and outcome of disease. As this thesis focuses on prognostic research, a detailed literature review of prognostic indicators and multivariate prognostic tools developed for breast cancer patients is conducted here.

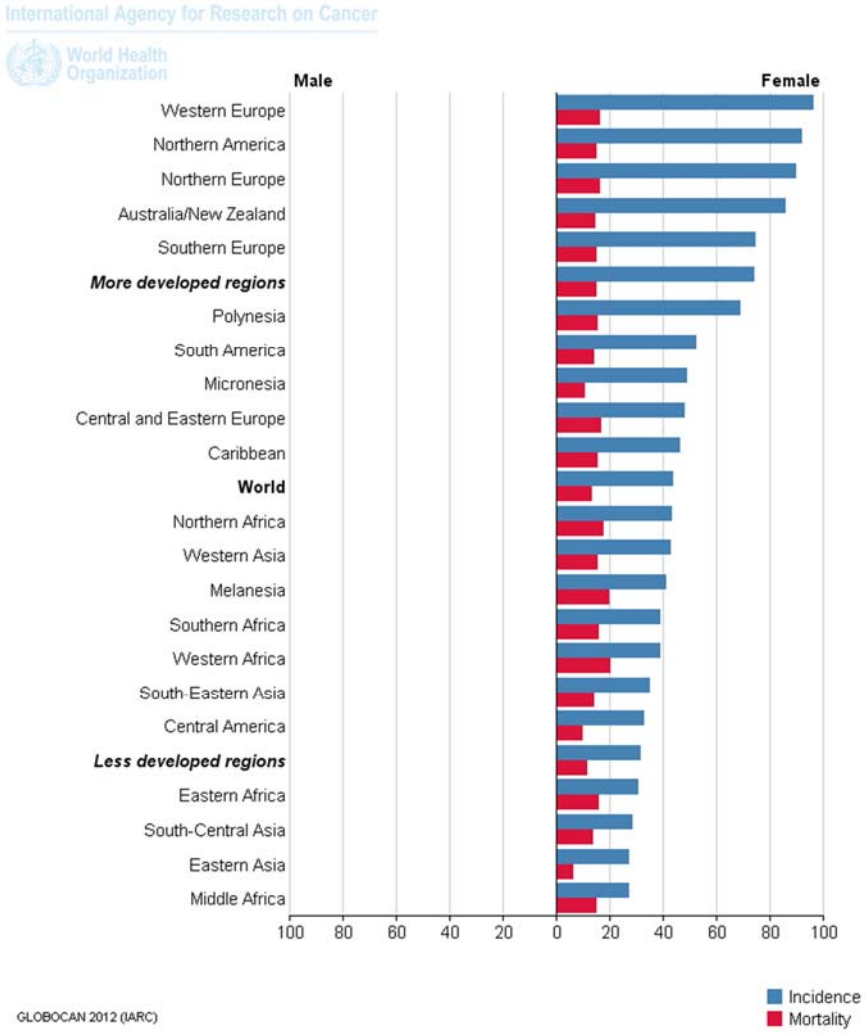
1.1 Burden of the disease

Breast cancer is the most common cancer among women in the world and is also one of the leading causes of cancer death among women according to GLOBOCAN 2012 statistics published by the International Agency for Research on Cancer (IARC) of the World Health Organization (WHO). In 2012, there were 1.7 million women diagnosed with breast cancer worldwide, which accounted for 11.9% of all new cancer cases that year[1]. The incidence of breast cancer is much higher in more developed regions comparing to less developed regions (Figure 1.1). It varies from 27 per 100,000 in Middle Africa and Eastern Asia to 96 per 100,000 in Western Europe [4].

Incidence has been rising in many parts of the world and the increment in Asia is more rapid than in the West, especially over the last decade (Figure 1.2) [5, 6]. There is a steady decline in mortality in North American and European countries (Figure 1.3). In contrast, mortality in Asian and Latin American countries continues to increase. As a result, the global mortality increased by 14% between 2008 and 2012. Models developed by the Cancer Intervention and Surveillance Modeling Network have demonstrated that decline in breast cancer death rate from 1975 to

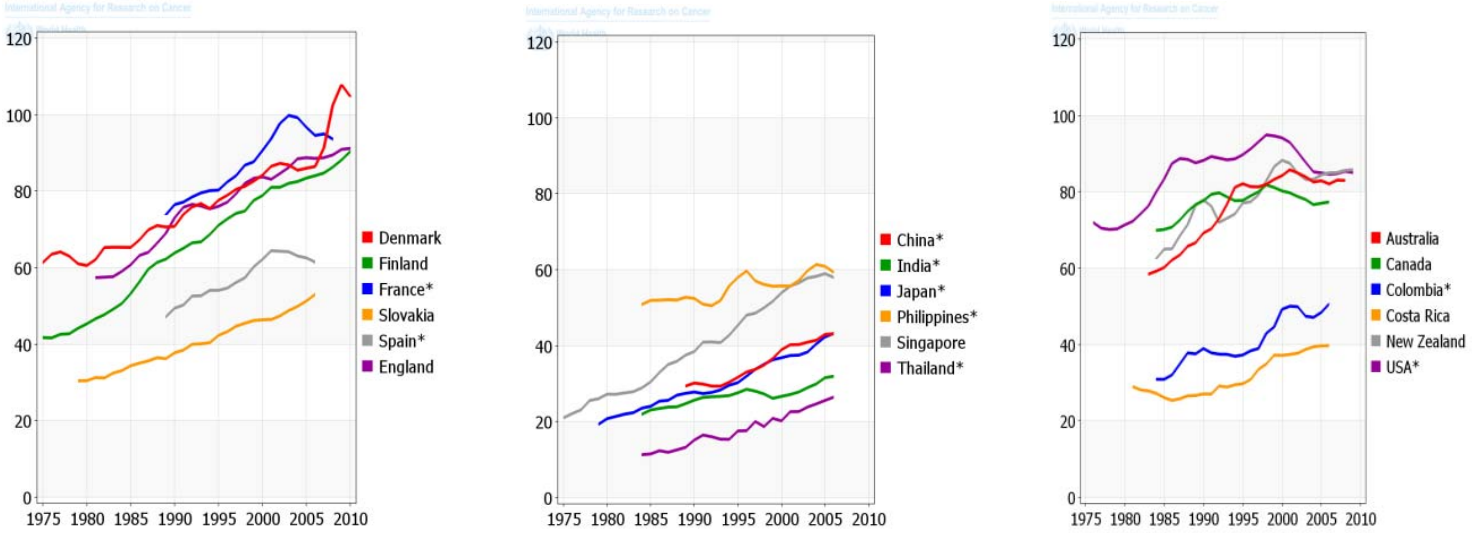
2000 in the United States is attributed to both mammography screening and adjuvant therapy [7]. This may imply huge inequality in terms of early detection and access to healthcare between rich and poor countries.

Figure 1.1 Estimated age-standardized rates (world population) per 100,000



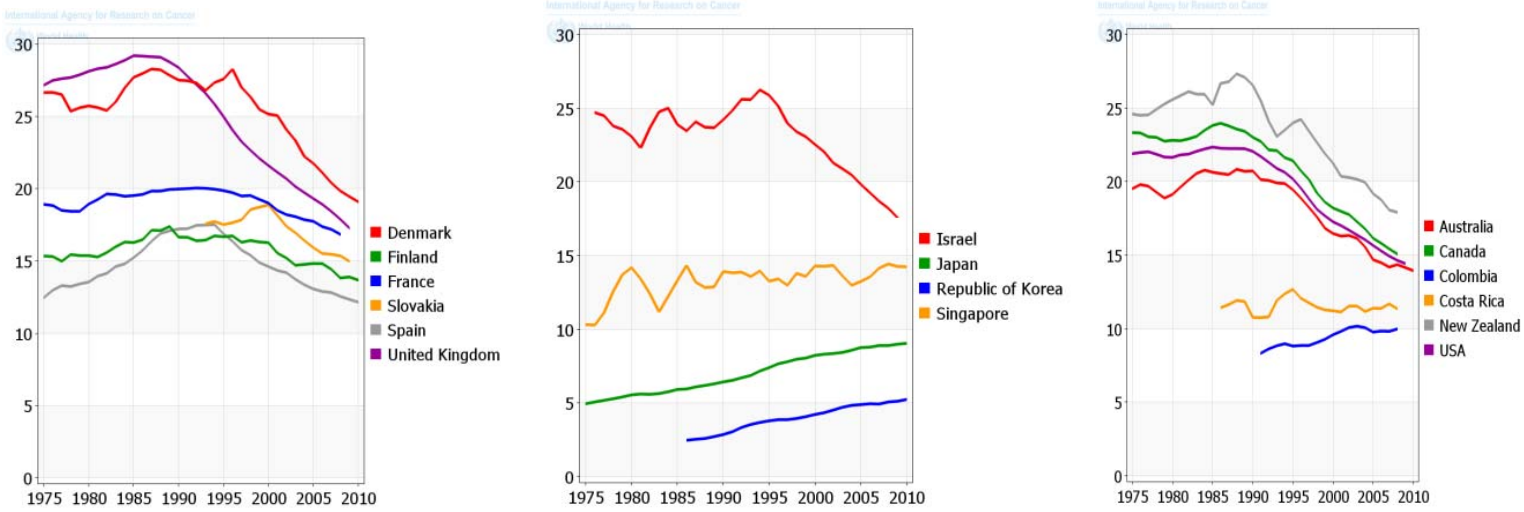
Source: Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013. Available from: http://globocan.iarc.fr, accessed on 1/12/2014.

Figure 1.2 Trends in incidence of female breast cancer in selected countries: age-standardized rate (world population) per 100,000



Source: Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013. Available from: http://globocan.iarc.fr, accessed on 1/12/2014.

Figure 1.3 Trends in mortality of female breast cancer in selected countries: age-standardized rate (world population) per 100,000



Source: Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013. Available from: http://globocan.iarc.fr, accessed on 1/12/2014.

Although the incidence in Asia is still much lower than that in the Western populations, it is increasing rapidly, most likely due to the adoption of Western lifestyles, reproductive patterns and access to early detection [8]. Singapore has one of the highest breast cancer incidence rates in Asia [5]. According to the report published by the National Registry of Disease Office (NRDO) in 2012, the incidence has increased nearly threefold from 21.5 per 100,000 in 1971-1975 to 60.7 per 100,000 in 2006-2010 [9]. One in 16 Singaporean women will develop breast cancer by age 75. The age-standardized mortality rate in Singapore increased from 8.5 per 100,000 in 1971-1975 to 13.5 per 100,000 in 1991-1995, but remained flat since then [9]. A study has estimated the period and cohort effects on breast cancer incidence in Singapore and Sweden, and found that cohort effect was much greater than period effect in Singapore, and was also more than what was observed in Sweden [8]. This finding implies that gradual change towards a more Westernized society in Singapore has contributed to the increasing incidence rate, especially for the more recent birth cohorts.

Breast cancer in men is very rare, and only accounts for less than 1% of all breast cancer cases in Europe and the United States. The incidence remained stable over the last few decades in Europe but increased recently in United States [10, 11] . Countries with higher incidence among women also have higher incidence among men [12].

1.2 Risk factors for breast cancer

Age

The risk of breast cancer increases with age. The incidence rises rapidly until the age of 50 years, which corresponds to average age of menopause, and continues to increase at a slower rate. The point of inflection in the age-specific incidence curve is known as the Clemmesen's hook. This pattern may be caused by the diminishing level of circulating estrogens after menopause [13]. Studies have suggested that breast cancer diagnosed in young (premenopausal) women is etiologically and biologically different from those in older (postmenopausal) women [14-17]. The effect sizes of several risk factors such as family history, obesity, weight gain and circulating endogenous estrogen vary between pre- and postmenopausal women [15, 18-21]. Tumors in young women are more likely to be of large size, high grade and hormone receptor negative [22, 23].

Comparing to Caucasian women, a higher proportion of Asian women are diagnosed at a younger age and premenopausal [5]. The peak age of onset in Asia is 45–50 years, whereas it is 60–70 years in the Western countries [24]. A study has found that the incidence rate of breast cancer among women younger than 45 years in Singapore become almost similar to that of Swedish women of similar age. In contrast, older women in Singapore had much lower incidence rate than their Swedish counterparts [8]. In the United States, a bimodal distribution of onset age was observed, with an earlier peak at 50 years and a smaller mode at 70 years [25].

Ethnicity

The risk of breast cancer differs among ethnic groups due to differences in exposure of established risk factors and genetic profile. In the United States, non-Hispanic White women have the highest incidence rate, while Asian American and Pacific Islander women have the lowest incidence [26]. In Singapore, incidence differs among the three major ethnic groups. The age-standardized incidence rate in Singapore from 2006 to 2010 was 64.3 per 100,000, 58.7 per 100,000 and 61.4 per 100,000 for Chinese, Malay and Indian women, respectively [9].

Family history

Approximately 15% of all breast cancer cases have a family history of breast cancer in first-degree relatives [27]. Women with a sister, a daughter or her mother diagnosed with breast cancer were twice as likely to develop breast cancer [28]. The risk was much higher if multiple first-degree relatives were affected or a relative was diagnosed at a young age. Besides shared environmental risk factors, studies which compared the monozygotic and dizygotic twins have suggested that 27% of the total risk of breast cancer can be explained by inherited genetic components [29]. High-penetrant germline mutations in *BRCA1* and *BRCA2* genes may account for 20% to 25% of inherited breast cancer cases and increase the lifetime risk of developing breast cancer to 40%-80% [30]. Relatively rare moderate-penetrant variants such as *PALB2*, *ATM* and *CHEK* may increase the lifetime risk of breast cancer by two-fold to 20% -50% [30]. A recent study published in the New England Journal of Medicine has concluded five to nine times increase in breast cancer risk to 33% -58% among *PALB2* mutation carriers [31]. To date, 72 common low-penetrant breast cancer susceptibility variants, which are

found in more than 5% of in the population have been identified by genome-wide association studies (GWAS) [30]. The effect sizes of these common alleles are generally modest with odds ratios less than 1.5 [32]. If we assume the combined effect of these loci is multiplicative, the top 5% of female population with the highest genetic risk based on genetic profile have an approximately 2.3-fold higher risk than the average population [33]. All these identified low to high penetrant loci together explain approximately one-third of the familial risk of breast cancer [33].

Mammographic density

Mammographic density measures the relative area of epithelial and connective tissue in the breast, which is radiographically dense and appears white on a mammogram, while the area of fat tissue appears translucent. Percent density, which is the dense area as a fraction of total breast area, is positively and almost linearly associated with breast cancer risk. Women with high percent density (more than 75%) are four to five times more likely to get breast cancer than women with low density (less than 10%) [34]. And every one standard deviation increment in percent density will increased the breast cancer relative risk by 52% for premenopausal women and 53% for postmenopausal women [35]. The association is independent of other risk factors such as age and parity status, and mammographic density can predict breast cancer risk for over 10 years. Comparing to percent density, absolute dense area is a weaker risk factor but still significantly associated with breast cancer risk after adjustment for confounders [35]. Mammographic density gradually decreases with age and large decline in percent density over time is associated with lower breast cancer risk [36, 37].

Reproductive factors

Breast cancer is a hormone-related disease. In particular estrogen has been shown to induce and promote mammary tumors. Certain reproductive factors that modify sex hormone levels, may also affect breast cancer risk. For example, the ovaries start to produce steroid hormone at around the time of menarche and gradually reduce their function at menopause. Women with earlier age at menarche or older age at menopause experience a longer period of increased level of circulating steroid hormone, and are thus more likely to develop breast cancer. Breast cancer relative risk increases by 5% for each year decrease in age at menarche, and 3% for each year that menopause is delayed [38]. Parous women have a lower risk than nulliparous women and the younger a woman has her first full-term birth the lower her risk of breast cancer [39]. Breast-feeding for an extended period reduces breast cancer risk in premenopausal women [40].

Hormone replacement therapy

Hormone replacement therapy (HRT) is prescribed to women who suffer from menopausal symptoms caused by diminished circulating estrogen and progesterone levels. The Women's Health Initiative study and the Million Women Study published in 2002 and 2003 reported that HRT was associated with a 1.2-to 2-fold increase of breast cancer risk, especially among current users of combined estrogen and progestogen HRT [41, 42]. Since then the use of HRT has dropped, and the magnitude of decline ranged from 34% to 79% in Western countries [43]. Consequently a few population-based studies also reported decline in breast cancer incidence rate between 2001 and 2006, ranging from 5% to 23% [43]. And such

decrease was only observed in women who were older than 50 years and more pronounced for estrogen receptor positive tumors [44].

1.3 Early detection and diagnosis of breast cancer

Physical examination

The most common signs and symptoms of breast cancer include a lump in a breast, change in appearance of the breast or nipple. During clinical breast examination, a health professional thoroughly palpates the breast and lymph nodes in the armpit and above the collarbones in a vertical strip pattern or a circular motion and inspects the breast and nipple for any changes of texture, shape and size. If examination findings are suspicious of breast cancer, more tests will be performed to confirm the diagnosis. Although clinical examination may detect cancer that is missed by mammography, there is no randomized controlled trial (RCT) to assess its impact on mortality reduction [45-47]. Breast self-examination does not significantly reduce breast cancer mortality based on meta-analysis of two large RCTs [47].

Imaging tests

Many women with early breast cancer do not have any symptoms. Therefore screening mammograms were introduced in many parts of world since 1980s, aiming to detect asymptomatic cancer at an early stage, thereby providing the opportunity for early intervention and better chance of survival. In most of countries with nationwide or regional screening programmes, women aged 40-69 are invited to screen from once a year to once every three years [48]. The participation rates vary between countries, from 18% in Japan to 87% in Finland [48]. The cost-effectiveness of a population-based screening programme by

mammography in Asia is debatable [49-51]. Thus far only high income countries in Asia such as Japan, Korea and Singapore have implemented such programmes. Opportunistic mammography screening by physician referral or self-referral is available in other Asian countries such as Hong Kong and Malaysia [52, 53]. In some European countries, both screening approaches co-exist and opportunistic screening is less sensitive and less cost-effective than organized screening programme [54-56].

The most reliable evidence on the effect of mammographic screening comes from RCTs. There have been 11 RCTs conducted in the 1970s and 1980s (Table 1.1), and a meta-analysis of 8 RCTs reported a 20% relative reduction in mortality for women who attended screening [57]. Another meta-analysis also found a relative risk of breast cancer mortality of 0.81 for the screened group compared to the control group and the difference was statistically significant [58]. However subgroup analysis suggested that the effect was not significant among women aged 40-49 years.

Negative findings on mortality reduction, potential risk of over-diagnosis and false positive results have generated controversies and debates over mammographic screening [59-61]. Over-diagnosis refers to diagnosis of indolent cancers which will not cause any symptom or death if left undetected and untreated. False positive mammograms lead to additional diagnostic tests, increase anxiety and decrease future screening participation [62-65]. Moreover screening misses interval cancers, which present between routine screens and are more aggressive and lethal [59]. The most recent publication from the Canadian National Breast Screening Study with 25-year follow-up showed no mortality reduction and estimated 22% of screen detected cancers to be over-diagnosed [61]. In the mammography arm of this trial,

27% of cancers were interval cancers and the survival of patients with interval cancers was much worse compared to patients with screen detected cancers [61]. An evaluation of 3 RCTs (Malmö I, Canada I and Canada II trials in Table 1.1) with relatively long follow-up and without screening invitation to the control group at end of the trial estimated the over-diagnosis to be 10% to 12% [57].

Diagnostic mammograms are used to diagnose women with symptoms or abnormal results from screening mammogram. Compared to screening mammograms, diagnostic mammograms include views of the area of interest from multiple angles. Ultrasound and magnetic resonance imaging can be used along with mammography to provide additional information such as distinguishing a solid mass from a cyst and determining its actual size.

Table 1.1 Characteristics of RCTs of breast cancer screening

	New York HIP	Malmö I and II	Swedish Two County	Canada I and II	Stockholm	Göteborg	UK Age trial	Edinburgh
Start Year	1963	1976	1977	1980	1981	1982	1991	1978
Number of women	62,000	60,076	133,065	89,835	60,800	52,222	160,921	54,654
Age group	40-64	45-69 and 43-49	38-75	40-49 and 50-59	39-65	39-59	39-41	45-64
Invited group intervention	M+PE	M	M+SE	M+PE+SE	M	M	M	M+PE
Screening interval (months)	12	18-24	24-33	12	24-28	18	12	24
Duration of screening (years)	3	12	7	5	4	7	8	6
Attendance	65%	74%	85%	88%	82%	84%	81%	65%
Control group intervention	None	None	None	PE+SE	None	None	None	None

M: Mammography PE: Physical examination SE: self-examination

Source: Independent UK Panel on Breast Cancer Screening, The benefits and harms of breast cancer screening: an independent review. Lancet, 2012. 380(9855): p. 1778-86.

Biopsy

All lesions suspicious of breast cancer need tissue confirmation. Cells or tissue from the suspicious area can be obtained by fine needle aspiration cytology or core biopsy. The collected sample is examined by a (cyto)pathologist for presence of malignant cells. Fine needle aspiration which uses narrower gauge needle is less invasive and the result is available immediately. Core biopsy provides architectural information in diagnosing invasive cancer and obtains more tissue to allow test for receptor status. Both methods are well-established with high degree of sensitivity and specificity. The false positive rate is less than 1% and false negative rate is about 3%-24% for fine needle aspiration cytology and 1%-2% for core biopsy [66]. Ultrasound, mammography or magnetic resonance imaging is used to guide the needle into the correct area.

1.4 Clinical characteristics and prognostic factors

Breast cancer is a heterogeneous disease and can be classified based on the histological type, grade, stage and expression of molecular or genetic tumor markers. These factors can be used for prognostication and selection of proper treatment.

Histological type

Breast cancer can originate from different types of tissue within the breast. Nearly all breast cancers start from the glandular tissues of the breast, such as the ducts and lobules. If the cancerous cells have not invaded through the basement membrane, it is defined as carcinoma *in situ*. Otherwise, it is defined as invasive or infiltrating carcinoma. Ductal carcinoma begins in the lining of a milk duct and is the most

common type of invasive and *in situ* breast cancer. About 70-73% of invasive diseases are invasive ductal carcinoma (IDC) [27]. Ductal carcinoma *in situ* (DCIS) is a pre-cancerous condition and has the potential to invade and become invasive cancer over time. Invasive lobular carcinoma (ILC) begins in the milk-producing lobules and is the second most common type of breast cancer, accounting for 13%-16% of invasive tumors. Lobular carcinoma in situ (LCIS) is not considered a cancer but a risk indicator for developing invasive breast cancer later on. Compared to IDC, ILC is less likely to present as a discrete mass, and therefore more difficult to be detected by mammography and palpation [67]. The difference in growth pattern can be explained by loss of E-cadherin expression frequently observed in ILC and LCIS tumors [68]. Expression microarray analysis has also identified a few genes expressed differently between ductal and lobular tumors. Although ILC is usually diagnosed at a later stage, the prognosis of ILC patients is better than IDC patients after adjustment for stage as ILCs are more likely to be hormone receptor positive [69]. However the prognostic advantage of ILCs disappears at approximately 6 years after diagnosis [70]. IDC and ILC also have distinct metastatic pattern. IDC is more likely to spread to lung and brain, while ILC commonly spreads to the gastrointestinal tract, gynaecological organs and peritoneum [71, 72].

Besides ILC and IDC, the remaining approximately 15% of invasive tumors comprise special types of breast carcinoma such as tubular, mucinous, medullary and papillary carcinoma. The cancer cells of these special types have different sizes, shape or growth pattern and each subtype accounts for less than 4% of all breast cancers [73, 74]. Rarely breast cancer can begin in the connective and support tissue of the breast, which is known as breast sarcoma.

Grade

The grade of tumor describes how different the cancer cells are from normal cells in size and shape and how fast the cancer cells are growing when viewed under a microscope. Tumor grade is commonly evaluated according to the Elston-Ellis modification of Scarff-Bloom-Richardson grading system (Nottingham Combined Histologic Grade) [75, 76]. In this system, percentage of glandular/tubular formation, the variation in size and shape of the nuclei (nuclear pleomorphism) and the number of dividing cells (mitotic activity) are taken into consideration and each element is given a score of 1 to 3 by the pathologist. The total score from the three components is further classified into three levels. Grade 1 or low grade tumor indicates well-differentiated cancer cells that are similar to normal cells and slow-growing, grade 2 or intermediate grade tumor composes of moderately-differently cancer cells and poorly-differentiated cancer cells which are completely different from normal cells, and fast-growing are categorized as grade 3 or high grade. Higher grade tumors grow and spread more rapidly, and are associated with increased risk of distant recurrent and poorer survival [77]. This association is independent of other predictors for survival such as tumor size and lymph node involvement [77].

Stage

The stage of breast cancer describes the extent to which the disease has spread and is determined by size of the tumor (T), number of lymph nodes carrying metastases (N) and presence of distant metastases (M). An overall stage 0-IV can be assigned to breast tumor when T, N and M are combined. This TNM staging system was first introduced by Union for International Cancer Control and then adopted by

American Joint Committee on Cancer (AJCC). The first edition of AJCC *Manual for Staging of Cancer* was published in 1977. Since then several changes have been made over time and the latest seventh edition was published in 2011. Some of the major changes from sixth edition to seventh edition included classification of small tumors with exclusively micrometastasis in lymph nodes as stage IB and creation of M0 (i+) category. Patients with metastasis in supraclavicular lymph node were considered N3 before 1987 and were re-classified as M1 in the fifth edition of AJCC staging system. The sixth edition, published in 2002, revised the classification of isolated ipsilateral supraclavicular metastasis from M1 (stage IV) to N3c (stage IIIC). Several studies have confirmed that survival of this group is more similar to locally advanced breast cancer rather than breast cancer with distant metastasis [78-80]. TNM staging can be based on physical examination, imaging test and laboratory results before initiation of treatment (clinical stage) or based on pathological examination of resected tumor and lymph nodes following surgery. Prefix "yc" and "yp" should be used for clinical and pathologic TNM stage evaluated after neoadjuvant treatment.

Tumor size and lymph node involvement are the strongest prognostic indicators for non-metastatic breast cancer and larger tumor size is also correlated with more involved lymph nodes [81]. The prognostic effect of tumor size is independent of nodal status and is shown for all stages including node-negative disease. A study found that 1cm decline in tumor size was associated with a 2.5% reduction in 15-year mortality and 1.5cm decline was associated with 10.8% reduction among node-negative patients. The impact was much greater for node-positive patients with 10.3% reduction for 1cm decline and 23% for 1.5cm decline [82]. Breast cancer patients with lymph node involvement had four to eight times higher mortality than

node-negative patients, and the risk increased with increasing number of positive lymph nodes [81].

Many cancer registries such as the Surveillance, Epidemiology, and End Results (SEER) Program also use localized, regional and distant staging system, which is determined by whether the cancer is confined within the organ of origin (localized), has spread to surrounding organs or tissues or lymph nodes (regional) or has spread to distant tissues or organs or to distant lymph nodes (distant) [83]. According to 18 population-based cancer registries in the SEER database in the United States, 61% of invasive cases were diagnosed at localized stage, 32% and 5% were diagnosed at regional and distant stage respectively in 2004-2010. Due to introduction of screening programme and increased awareness of breast cancer, more breast cancers are detected at an early stage. However women in developing countries and women with lower socioeconomic status are still more likely to be diagnosed at later stages [84].

Survival of breast cancer patients varies by stage. The 5-year and 10-year survival of stage 0 (i.e. DCIS) patients is 99% and 98% respectively while stage IV (i.e. metastatic) breast cancer patients have 5-year survival of less than 25% and 10-year survival of approximately 7%.

Receptor status and molecular subtypes

Hormone receptors are proteins, to which specific hormone such as estrogen and progesterone will bind, thereby promoting the growth of cells. For breast cancer, the presence of estrogen receptor (ER) and progesterone receptor (PR) will affect the prognosis and treatment of cancer. ER positive breast cancer, which depends on estrogen to grow, can be treated with hormone therapy such as tamoxifen and

aromatase inhibitors or ovarian suppression or ablation for pre-menopausal women, aiming to prevent estrogen from binding to receptor or to reduce the production of estrogen. A meta-analysis of RCTs confirmed that tamoxifen for 5 years reduces recurrence and mortality risk for ER+ but not for ER- breast cancer [85]. Even without adjuvant systemic treatment, patients with ER+ tumor have been shown to have better disease-free survival and overall survival in NSABP B-06 trial [86]. The presence of PR is strongly correlated with ER status and is an independent predictor for benefit from adjuvant hormone therapy among ER+ patients [87]. In the SEER database, 63% of patients were ER+/PR+ and the remaining cases were ER+/PR- (13%), ER-/PR+ (3%) and ER-/PR- (21%) [88]. Comparing to ER+/PR+ tumor, women with ER+/PR-, ER-/PR+ and ER-/PR- tumor had 1.4-, 1.8- and 2.3-fold increased relative risk of death respectively [88]. Hormone receptor status is evaluated by pathologists, using immunohistochemistry (IHC) to determine percentage of stained tumor cells. Historically, a cut-off value of 10% was used to define ER or PR positivity. However, a few studies reported treatment response among patients with as low as 1% of cells stained positive [89]. In 2010, the American Society of Clinical Oncology and College of American Pathologists recommended 1% threshold to allow more patients eligible for hormone therapy [89].

Overexpression of human epidermal growth factor 2 (HER2) protein is associated with more aggressive tumor behavior, higher risk of recurrence and poorer survival, especially among node positive patients [90, 91]. It is caused by amplification of the *HER2* gene which can be found in 15-20% of patients with breast cancer [92, 93]. It can be assessed by either IHC to check for over-expression of HER2 protein or fluorescence *in situ* hybridization (FISH) to determine the amplification of *HER2*

gene. IHC test gives a score of 0 to 3+, based on staining intensity. The score can be interpreted as HER2 positive (3+), HER2 negative (0 or 1+) or equivocal (2+). FISH test provides the copy number of *HER2* gene per nucleus or the ratio of *HER2* to chromosome enumeration probe 17 (CEP17). According to Food and Drug Administration (FDA), cut-off value of four copies per nucleus or *HER2*/CEP17 ratio of 2 is used to define HER2 positivity [94]. HER2 overexpression used to be associated with a worse outcome. It is a predictive biomarker for response to chemotherapy, hormone therapy, and most importantly for HER2-targeted therapy such as trastuzumab and lapatinib. Use of trastuzumab has been shown to significantly improve overall survival and disease-free survival in HER2 positive breast cancer at all stages [95, 96]. Today, HER2 positivity is associated with improved outcome if treated with targeted therapy. The American Society of Clinical Oncology and College of American Pathologists recommend testing all newly diagnosed invasive breast cancers for HER2 status using either IHC or FISH [97].

Using DNA microarray analysis, researchers have found that breast cancer can be categorized into several subgroups based on gene expression profile of the tumor. The five major subtypes are luminal A, luminal B, HER2-overexpressing, basal-like and normal-like and they are associated with different outcomes. Joint ER, PR and HER2 status, and expression of three other biomarkers Ki-67, cytokeratin (CK) 5/6 and epidermal growth factor receptor (EGFR) are often used as surrogates of molecular subtypes as shown in Table 1.2 [27, 98, 99]. Triple negative breast cancers defined by absence of ER, PR and HER2 are more aggressive and have higher risk of recurrence and death due to lack of targeted therapy [100]. It is more

prevalent among premenopausal women and among women of African and South Asian descent than non-Hispanic whites [101-103].

Table 1.2 Characteristics of breast cancer subtypes defined by gene expression profile

	Luminal A	Luminal B		HER2 over-expressing	Triple negative	
		HER2 negative	HER2 positive		Basal-like	Normal-like
Biomarkers expression pattern	ER+ and/or PR+ HER2- Ki-67-	ER+ and/or PR+ HER2- Ki-67+	ER+ and/or PR+ HER2+	ER- PR- HER2+	ER- PR- HER2- CK5/6+ and/or EGFR+	ER- PR- HER2- CK5/6- EGFR-
Distribution	55-65%	7-12%		6-10%	10-15%	5-10%
Grade						
1 and 2	42%	44%		30%	18%	19%
3	58%	56%		70%	82%	81%
Stage						
I	44%	39%		28%	24%	48%
II	47%	54%		53%	62%	39%
III-IV	9%	6%		19%	13%	13%
5-year survival	75% - 90%	45% -90%		20% -75%	30% - 80%	50%- 87%

Source: Phipps, A.I. and C.I. Li, Breast Cancer Biology and Clinical Characteristics, in Breast Cancer Epidemiology, C.I. Li, Editor. 2010, Springer Science+Business Media.

Multiple genomic assays such as Oncotype DX and MammaPrint have been developed to estimate risk of recurrence for early breast cancer [104]. MammaPrint was developed based on expression of 70 genes within a tumor and can identify women in whom chemotherapy can be safely omitted [105]. OncotypeDX, which was developed based on 21 genes, can be used to predict benefit from hormone- and chemotherapy for ER-positive breast cancer [106].

Age

Many studies suggested that breast cancer diagnosed in young women (less than 35 years old) is biologically more aggressive (high grade, ER- and HER2+ tumors) than that in older women and young women are often diagnosed at a later stage. As a result prognosis is much poorer in terms of both survival and recurrence [107-111]. The negative prognostic effect of young age was also observed among premenopausal breast cancer [112, 113], while among postmenopausal women, older age was associated with higher mortality risk in stage I and II disease but not in stage III and IV disease, after adjustment for comorbidities [114]. Disease-specific mortality also increased with increasing age after adjustment for treatment and tumor characteristics in a study on postmenopausal hormone receptor positive breast cancer [115]. Two population-based studies in Singapore and Sweden found the highest relative survival among women to be between 40 and 49 years old [116, 117]. A Swedish study investigated the contribution of various determinants to the survival discrepancy between different age groups and concluded that tumor characteristics rather that treatment activity was the most important explanatory variable [118].

Comorbidity

Comorbid chronic conditions such as hypertension, diabetes and heart disease are often present at time of diagnosis of breast cancer, especially among elderly patients. In the SEER database, the most common comorbidities among women aged 65 and above were previous cancer (16%), diabetes (13%), chronic obstructive pulmonary disease (9%) and congestive heart failure (7%) [119]. Patients with comorbidities were less likely to receive radiotherapy or have axillary dissection if they underwent

lumpectomy in a Dutch study [120]. Many studies have reported that comorbidities are associated with poorer prognosis among breast cancer patients regardless of age and stage when they are measured using index or a sum of the number of conditions [121, 122]. However the effect of each individual type of comorbidity varies and whether it affects breast-specific mortality is uncertain. For example, diabetes was found to be significantly associated poor total and non-breast cancer mortality, but its effect on recurrence or breast cancer-specific mortality was not consistent across studies [123-125]. Studies also found liver disease, chronic renal failure, dementia, and congestive heart failure were the most significant conditions for overall mortality [119, 120].

Gender

Men are more likely to be diagnosed with breast cancer at an older age and more advanced stages as compared to women [126, 127]. However there are conflicting findings whether male and female patients have comparable survival after adjustment for age and stage. An important confounding factor would be difference in underlying life expectancy between men and women. Two population-based studies reported that although the overall survival of male breast cancer is poor for all stages, survival differences disappear after adjustment for background mortality rate, especially for early stage breast cancer [11, 126].

1.5 Treatment for breast cancer

Locoregional therapy

The goal of locoregional therapy for early breast cancer is to eradicate all malignant breast tissue in the breast and lymph nodes to achieve optimal local control with

surgery and radiotherapy [128]. The Early Breast Cancer Trialists' Collaborative Group has suggested that adequate local treatment can reduce 15-year breast cancer mortality by 5.2% by avoiding local recurrence [129]. Surgical removal of the primary tumor in breast includes either mastectomy or breast-conserving surgery. Whole breast irradiation is routinely given after breast-conserving surgery. Several RCTs with 20-year follow up have showed that long term distant-disease-free survival and overall survival among women who undergo breast-conserving surgery followed by irradiation is the same as that among women who undergo radical or total mastectomy [130-132]. However the incidence of recurrence in the ipsilateral breast is significantly higher in women treated with breast-conserving surgery than in those treated with mastectomy. Radiotherapy after breast-conserving surgery reduces locoregional and distant recurrence, and breast cancer death [133]. Most patients with a unifocal tumor and adequate tumor to breast size ratio are considered suitable for breast-conserving surgery [134]. For patients with multicentric disease and those who are unlikely to achieve negative surgical margins, mastectomy is the treatment of choice. Post-mastectomy radiotherapy may reduce recurrence and mortality in patients with high risk disease (i.e. large tumors, high nodal burden) [129, 135].

Axillary lymph node dissection (ALND) was the standard surgical procedure to remove lymph nodes in the axilla to evaluate the extent of disease (stage) [136, 137]. The survival benefit of ALND for clinical node negative patients remained controversial, given the evidence against ALND from NSABP B-04 trial and evidence from a meta-analysis of 6 RCTs which showed improvement in overall survival [138, 139]. Recent RCTs have reported that ALND, which is associated with complications such as lymphedema, pain and seroma, is unnecessary if cancer

is not present in the sentinel lymph node (SLN) [140-142]. SLN is defined as the hypothetical first lymph node(s) to which cancer cells will spread from a primary tumor and can be located by injection of radioactive substance or/and a blue dye near the tumor. SLN dissection has been widely accepted as staging tool due to lower morbidity. The Z0011 trial conducted by the American College of Surgeons Oncology Group found that ALND did not improve overall or disease-free survival even for positive SLN patients with T1-T2 tumor and treated with lumpectomy, whole breast irradiation and adjuvant therapy [143]. These findings have changed current practice in many parts of Europe.

For metastatic breast cancer, resection of the primary tumor is not recommended except for patients with symptomatic local breast or chest wall tumors [144]. Recent observational studies have suggested that surgical treatment of primary tumor could improve survival of patients with *de novo* metastatic breast cancer [145]. However it is difficult to sufficiently control for potential confounding by indication in these studies. The preliminary result from RCT conducted at Tata Memorial Hospital in Mumbai, India did not show any survival benefit of surgery in patients with *de novo* metastatic breast cancer [146].

Systemic therapy

Unlike local therapy which focuses on the primary tumor, systemic therapy aims to eliminate micrometastasis or microscopic tumor cells that have spread beyond the breast and nearby lymph nodes to the entire body. Systemic therapy consists of chemotherapy, hormonal therapy, and targeted therapy and can be given before or after locoregional treament (called neoadjuvant and adjuvant therapy, respectively). For early breast cancer, adjuvant therapy is more likely to be administered to

patients with higher risk of relapse or death, which can be determined based on prognostic factors discussed earlier (age, stage, receptor status etc.). Adjuvant tamoxifen, a hormone therapy drug for hormone receptor positive breast cancer reduces the relative risk of breast cancer death by 31% for ER+ breast cancer [147]. Five-year use is significantly more effective than use for 1-2 years [147]. Aromatase inhibitors (AI), another hormone therapy drug, gained popularity among postmenopausal women due to its greater effect in recurrence reduction [148]. Survival benefit of hormonal treatment is seen in the absence of chemotherapy and is still present, although less substantial when combined with chemotherapy. For adjuvant chemotherapy, multiple cytotoxic drugs are given in combination 3-4 weekly for four to six cycles. The Early Breast Cancer Trialists' Collaborative Group review published in 2005 has shown that second generation anthracycline-based polychemotherapy reduced annual breast cancer mortality by 38% for women younger than age 50 and 20% for women aged 50-69, and was more effective than first generation alkylating-based CMF (cyclophosphamide, methotrexate, fluorouracil) chemotherapy [147]. A review published in 2012 confirmed that the third generation taxane-plus-anthracycline-based regimens slightly but significantly improved outcome in comparison with an anthracycline-based control regimen [149]. And the proportional reduction in recurrence and breast cancer mortality due to taxane-based regimens or anthracycline-based regimens was independent of age, nodal status, tumor size, grade and ER status [149]. Adjuvant trastuzumab, the first targeted therapy for HER2+ cancer, was introduced to early breast cancer after robust result were observed in metastatic breast cancer setting. A meta-analysis of 5 RCTs published in 2008 demonstrated lower mortality and

recurrence risk among non-metastatic patients who received adjuvant trastuzuma and chemotherapy versus patients who received chemotherapy alone[150].

Systemic therapy can also be given to locally advanced disease before surgery to make inoperable tumor operable. For early stage breast cancer, neoadjuvant systemic treatment increases the likelihood of receiving breast-conserving surgery instead of mastectomy by shrinkage of tumor. Survival is not affected by timing of chemotherapy (i.e. neoadjuvant versus adjuvant) [151]. For patients with hormone receptor positive tumors and are unfit for chemotherapy, neoadjuvant hormone therapy can be considered [152]. For patients with HER2+ tumors, addition of targeted therapy has shown better result than chemotherapy alone [153]. Pathological complete response (pCR) after neoadjuvant treatment is strongly associated with disease-free survival and overall survival. Therefore it is commonly used as an intermediate endpoint in clinical trials [154]. The three most common definitions of pCR are 1) absence of invasive cancer and *in-situ* cancer in the breast and axillary nodes, 2) absence of invasive cancer in the breast and axillary nodes, irrespective of DCIS, and 3) absence of invasive cancer in the breast irrespective of DCIS or nodal involvement [155]. The prognostic value of pCR is dependent of receptor status and grade, as the strongest association was found in the aggressive subtype [155, 156].

For patients with metastatic breast cancer at presentation, the goal is to palliate symptoms and prolong life since cure is highly unlikely. For these patients, systemic therapy is the mainstay of treatment. Hormone therapy is the preferred initial treatment for ER+ and/or PR+ tumor unless there are symptomatic metastases at visceral organs, for which chemotherapy will be given. Unlike early breast cancer, single agent chemotherapy is preferred over polychemotherapy

regimen for most metastatic patients except for those with rapid progression or life-threatening visceral metastases [157]. For recurrence at distant sites, there is high chance of drug resistance to previously administered chemotherapy or hormone therapy.

Adverse effects of treatment

Although tremendous advances in systemic treatment have contributed to considerable improvement in breast cancer outcome, many patients experience short-term and long-term adverse effects caused by the treatments. Common short-term complications include nausea and vomiting, mucositis, neutropenia, myelosuppression and fatigue, which occur during the course of treatment [158]. Long-term side effects such as cardiac toxicity, bone fracture, secondary cancer, cognitive impairment, menopausal symptoms and infertility may impair the patients' quality of life, years or even decades after completion of treatment [159, 160]. Breast surgery, ALND and radiotherapy may also cause seroma, fibrosis, lymphedema, and cardiovascular disease [161].

1.6 Breast cancer prognosis

Locoregional recurrence and distant relapse

Breast cancer can relapse after initial treatment, impairing survival. The most common sites of relapse are listed in Table 1.3. The suspicion of relapse is confirmed by imaging tests such as ultrasound, computed tomography, magnetic resonance imaging and positron emission tomography. Studies conducted before the widespread use of breast-conserving surgery and systemic therapy have shown

that distant metastasis to skeleton is the most frequent site of first relapse, followed by locoregional recurrence and pulmonary metastasis [162]. Pattern of spread can differ by histological types and receptor status. IDC is more likely to spread to lungs and brain and ILC commonly spreads to the gastrointestinal tract, gynaecological organs and peritoneum [163]. Distant metastasis of ER- disease is more likely to be seen in soft tissues such as lungs, liver and brain and patients with HER2 positive tumors have an increased risk of brain metastasis [164]. In the study of recurrence in patients enrolled in seven Eastern Cooperative Oncology Group trials, 45% experienced recurrence with a peak incidence between 1 and 2 years after diagnosis, followed by a decrease [165]. The rate of decrease was much slower from five years onwards as compared to from 2nd to 5th year. For breast carcinoma *in situ*, the risk of subsequent ipsilateral invasive breast cancer is higher while the risk of distant metastasis is still low [162].

Table 1.3 Most common sites of breast cancer relapse.

Type of relapse	Location
Local recurrence	Ipsilateral breast (breast conserving surgery) or chest wall (mastectomy) including skin and subcutaneous tissue, surgical scar and biopsy tract
Regional recurrence	Ipsilateral lymph nodes (axillary, supraclavicular, infraclavicular, internal mammary, and intramammary).
Distant recurrence	Contralateral lymph nodes (axillary, supraclavicular, infraclavicular, parasternal, and internal mammary) in absence of synchronous ipsilateral or contralateral breast malignancy or distant metastasis, skin and subcutaneous tissue outside the ipsilateral chest wall, bone, lung, liver and central nervous system

Source: Moossdorff M et al, Maastricht delphi consensus on event definitions for classification of recurrence in breast cancer research. J Natl Cancer Inst. 2014 Nov 7;106(12). pii: dju288.

Second primary cancer

Breast cancer survivors have a 2- to 6-fold increased risk of developing a second cancer in the contralateral breast as compared to healthy women developing a first breast cancer [166]. Their risk of developing a second non-breast primary cancer is 25% higher [167]. For DCIS patients, their risk of developing subsequent invasive or *in situ* breast cancer in the contralateral breast is also higher than the general population [168, 169]. Risk factors for contralateral breast cancer include family history, age at first diagnosis and ER status of first tumor [170]. Women with bilateral breast cancer have higher mortality compared to women with unilateral disease. The risk increases further if the time interval between the first and second cancer is shorter [171]. For other secondary cancers, the increased risk at sites such as esophagus, lungs, stomach and soft tissues might be consequences of radiation used to treat breast cancer [167]. Certain cytotoxic drugs for breast cancer have been shown to associate with higher risk of acute myeloid leukemia, and tamoxifen increases the risk of endometrial cancer [172]. Increased surveillance and shared genetic and environment risk factors may also play a role in the increased risk of secondary cancer [173].

Survival

The 5-year overall survival of breast cancer was 67.9% in Singapore in 2003-2007 [174]. Compared to the general population, the 5-year age-standardized relative survival (ASRS) was 76.4% during the same period [174]. The estimate in Singapore was slightly lower than high-income countries such as the United States, Sweden, Japan, Finland and Australia, where the ASRS was more than 80% [175]. The ASRSs in middle and low-income countries were less than 60% and less than

40% respectively [175]. Due to advances in early detection and treatment, relative survival in many countries has improved over time. In England and Wales, 1-year relative survival increased from 82% in 1971-1975 to 96% in 2005-2009 and 5-year relative survival increased from 52% to 85% [176, 177]. A similar trend has also been observed in Singapore, where 5-year ASRS increased from 49% in 1973-1977 to 76% in 2003-2007.

There are remarkable ethnic and socioeconomic disparities in survival after breast cancer [175, 178-180]. In the United States, African American patients have significantly worse survival than Caucasian women after controlling for differences in socioeconomic status [181]. Other ethnic groups such as Latinas and Pacific islanders have intermediate prognosis compared to African and white Americans, whereas survival in Asian Americans varies among different subpopulations and by immigrant status [182, 183]. In Singapore and Malaysia, Malay ethnicity is associated with poorer outcome compared to the other two major ethnic groups, namely, Chinese and Indians, after adjustment for age, stage and tumor characteristics [178]. The ethnic disparities in survival can be explained by differences in life expectancy, comorbidities, socioeconomic status, tumor biology, response to treatment, and lifestyle before and after diagnosis. A review on socioeconomic differences in cancer survival was conducted by IARC in 1997 and it showed that female breast cancer was one of the cancer sites with the widest differences, regardless of which measure of socioeconomic status was used [184]. Poorer survival of patients from lower socioeconomic status can be explained by poorer access to health care, leading to late presentation with more advanced stages at diagnosis, and suboptimal treatment.

Multivariate prognostic tool

As described in the previous sections, breast cancer is a complex and heterogeneous disease with various treatment options. Accurate prognostication is important for clinicians and for patients to benefit the most from adjuvant therapy and spare them from toxic side effects, reduce financial burden, and to let them have an idea about life expectancy. Many prognostic tools have been developed to estimate risk of recurrence and death as well as added benefit from specific type of treatment. These tools include predictors such as traditional clinicopathologic factors (age, tumor size and number of positive nodes etc.) and/or gene expression profile of the tumor and novel biomarkers. The selection of predictors and statistical/mathematical algorithms used to combine their effect size vary between different prognostic tools. A systematic review published in 2014 compared the characteristics of several risk prediction models for non-metastatic breast cancer (Table 1.4) [185].

The most widely used prognostic models are the Nottingham Prognostic Index (NPI), Adjuvant! Online, MammaPrint and OncotypeDx, and they have been validated in different populations. NPI, introduced in 1982, was the first prognostic model for breast cancer patients. It includes only tumor grade, size, and nodal status for prediction of disease-free survival, and can be calculated as 0.2×Size (cm)+ lymph node stage (1-3) + grade (1-3) [186, 187]. A higher score indicates worse survival and the score is usually split into three to five subgroups. The widely used Adjuvant! Online (www.adjuvantonline.com) (Figure 1.4) is a software model, that calculates 10-year overall survival and disease-free survival of patients with non-metastatic breast cancer, based on a patient's age, tumor size, grade, ER status, nodal status, and comorbidities. It also quantitatively predicts the absolute gain from adjuvant therapy [188]. CancerMath (http://www.lifemath.net/cancer/)

(Figure 1.5) is the latest web-based calculator, which takes HER2 status into account when estimating overall survival for each of the first 15 years after diagnosis [189]. It was established based on the binary biological model of cancer metastasis and the parameters of the mathematical equation were derived from the SEER database in the United States. It was shown to be accurate and comparable with Adjuvant! Online [189]. Most of these models were accurate in populations similar to its derivation dataset and were able to stratify patients into low, moderate and high risk groups [185]. Adjuvant! Online was associated with 0.9% to 12.7% change in treatment recommendation in four prospective studies [190]. Although it was recommended by the National Institute for Health and Clinical Excellence and widely used by oncologists [191-194], several validation studies suggested that Adjuvant! Online was suboptimal in women younger than 40 years and older than 75 years [195, 196]. The model was recently validated in Malaysia, Korea, and Taiwan, where it was shown to substantially overestimate actual survival [197-199]. The underperformance can be attributed to different underlying mortality risk, proportion of advanced disease, tumor biology, response to treatment, and lifestyle after diagnosis between US population and study populations [200-202].

Patients tend to misunderstand the results provided by these tools [203]. Studies have found that prognostic assessment may cause anxiety and distress among patients [203, 204]. Therefore clear and simplified illustration with graphic representation and explanation by trained health professionals are needed.

Table 1.4 Summary of risk prediction models

Name	Year	Country	Outcome	Predictors	Target patients	Forms/ platform
NPI [186]	1982	UK	RFS	Grade, T, N, LVI		Mathematical formula
Adjuvant ! Online [188]	1996	US	OS+RFS +Tx benefit	Age, grade, comorbidities, T, N, ER, Tx regimen		Web-based
BC Nomo-gram [205]	2004	US	BCSS	Age,histologic type, T, multifocality, LVI, staining		Nomogram
OPTIONS [206]	2010	UK	RFS	Age, grade, T, N, ER, Tx regimen		Web-based
PREDICT [207]	2010	UK	OS	Age, grade, T, N, ER, , Tx regimen (chemo only), mode of detection		Web-based
Cancer-Math [189]	2011	US	OS+Tx benefit	Age, histologic type, grade, size, T, N, ER, HER2, Tx regimen		Web-based
PREDICT Plus [208]	2012	UK	OS+ BCSS	Age, grade, T, N, ER, HER2, Tx regimen (chemo only), mode of detection		Web-based
Sigudsson et al [209]	1990	Sweden	RFS	T, PR, S-phase category	Node negative	
Aubele et al [210]	1995	Germany	RFS	T, N, Dra_CV	Node positive	
Mamma-Print [105]	2002	Netherla-nds	RFS	GP, 70 genes	T<5 cm N<=3	Micro-array
Oncotype Dx [106]	2004	US	RFS	GP, 21 genes	N0, ER+	RT-PCR
WR Signature [211]	2004	US	RFS	GP, 512 genes		Micro-array

Pawitan et al [212]	2005	Sweden	RFS	GP, 64 genes		Micro-array
Rotterdam Sig [213]	2005	Netherla-nds/US	RFS	GP, 76 genes	N0	Micro-array
Mammo-strat [214]	2006	Canada/ US	RFS	IHC, 5 proteins	ER+	
Ma et al [215]	2007	US	RFS	GP, 28 genes		Micro-array
Xu et al [216]	2008	US	RFS	GP, 112 genes		Micro-array
Theros BCI [217]	2008	US/UK/ Belgium	RFS	GP, 7 genes	N0, ER+	RT-PCR
PAM50 [218]	2009	US	RFS+Tx benefit	GP, 55 genes		RT-qPCR
GENIUS [219]	2010	Belgium	RFS	GP, 85 genes and AURKA	N0	Micro-array
Breast PRS [220]	2011	US	RFS	GP, 200 genes		Micro-array

BCCS, breast cancer specific survival; OS, overall survival; RFS, recurrence-free survival; Tx, treatment; GP, genetic profile; IHC, immunohistochemistry panel; Tm tumor size; N, nodal status; LVI, lymphovascular invasion

Source: Engelhardt, E.G., et al., Predicting and communicating the risk of recurrence and death in women with early-stage breast cancer: a systematic review of risk prediction models. J Clin Oncol, 2014. 32(3): p. 238-50

Figure 1.4 Screenshot of Adjuvant! Online

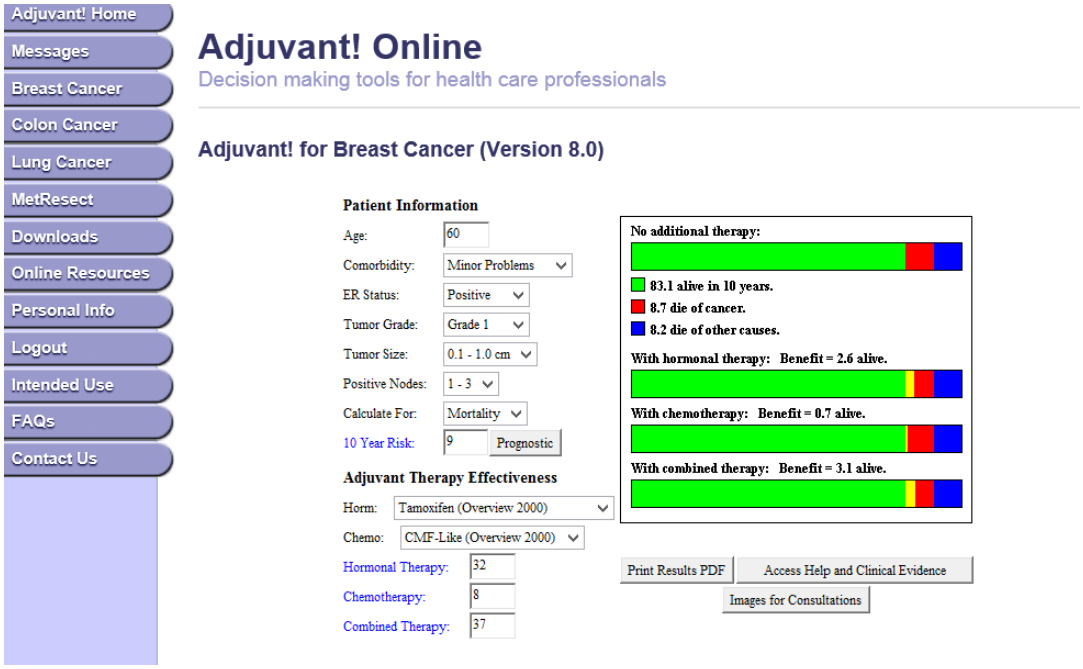
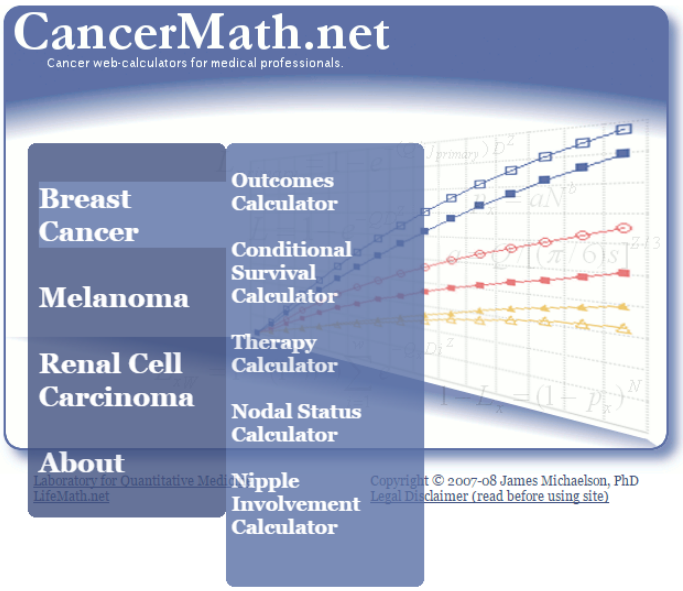
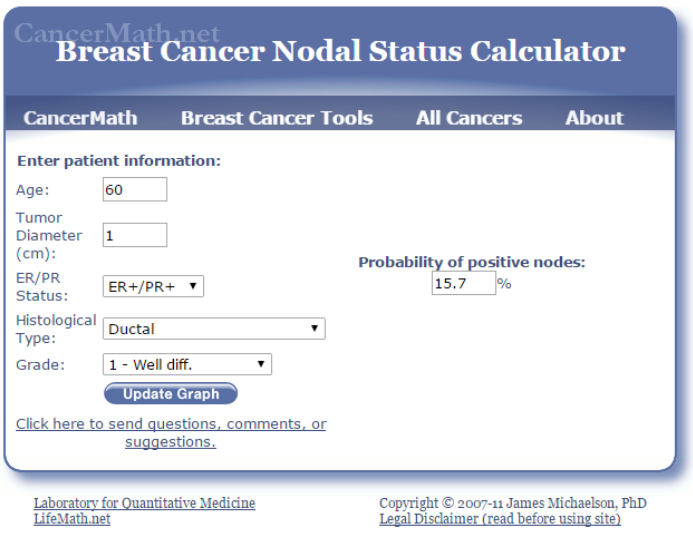
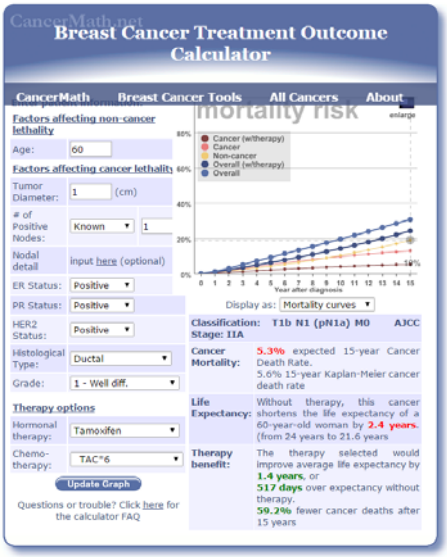


Figure 1.5 Screenshots of CancerMath.net




CancerMath.net
Breast Cancer Nodal Status Calculator
CancerMath
Breast Cancer Tools
All Cancers
About
Enter patient information:
Age:
60
Tumor Diameter (cm):
1
ER/PR Status:
ER+/PR+
Histological Type:
Ductal
Grade:
1 - Well diff.
Update Graph
Click here to send questions, comments, or suggestions.
Probability of positive nodes:
15.7
%
Laboratory for Quantitative Medicine
LifeMath.net
Copyright © 2007-11 James Michaelson, PhD
Legal Disclaimer (read before using site)


CancerMath.net
Breast Cancer Outcome Calculator
CancerMath
Breast Cancer Tools
All Cancers
About
Enter patient information:
Factors affecting non-cancer lethality
Age:
60
Factors affecting cancer lethality
Tumor Diameter:
1
(cm)
# of Positive Nodes:
Known
1
Nodal detail
input here (optional)
ER Status:
Positive
PR Status:
Positive
HER2 Status:
Positive
Histological Type:
Ductal
Grade:
1 - Well diff.
Update Graph
Questions or trouble? Click here for the calculator FAQ
mortality risk
enlarge
Cancer
Non-cancer
Overall
Year after diagnosis
Display as:
Mortality curves
Classification:
T1b N1 (pN1a) M0
AJCC Stage: IIA
Cancer Mortality:
11% expected 15-year cancer death rate.
(11.8% 15-year Kaplan-Meier cancer death rate)
Life Expectancy:
This cancer shortens the life expectancy of a 60-year-old woman by 2 years. (from 24 years to 22 years)
Laboratory for Quantitative Medicine
LifeMath.net
Copyright © 2007-11 James Michaelson, PhD
Legal Disclaimer (read before using site)


CancerMath.net
Breast Cancer Treatment Outcome Calculator
CancerMath
Breast Cancer Tools
All Cancers
About
Factors affecting non-cancer lethality
Age:
60
Factors affecting cancer lethality
Tumor Diameter:
1
(cm)
# of Positive Nodes:
Known
1
Nodal detail
input here (optional)
ER Status:
Positive
PR Status:
Positive
HER2 Status:
Positive
Histological Type:
Ductal
Grade:
1 - Well diff.
Therapy options
Hormonal therapy:
Tamoxifen
Chemo-therapy:
TAC*6
Update Graph
Questions or trouble? Click here for the calculator FAQ
mortality risk
enlarge
Cancer (w/therapy)
Cancer
Non-cancer
Overall (w/therapy)
Overall
Year after diagnosis
Display as:
Mortality curves
Classification:
T1b N1 (pN1a) M0
AJCC Stage: IIA
Cancer Mortality:
5.3% expected 15-year Cancer Death Rate.
5.6% 15-year Kaplan-Meier cancer death rate
Life Expectancy:
Without therapy, this cancer shortens the life expectancy of a 60-year-old woman by 2.4 years. (from 24 years to 21.6 years
Therapy benefit:
The therapy selected would improve average life expectancy by 1.4 years, or 517 days over expectancy without therapy.
59.2% fewer cancer deaths after 15 years
Laboratory for Quantitative Medicine
LifeMath.net
Copyright © 2007-11 James Michaelson, PhD
Legal Disclaimer (read before using site)


CancerMath.net
Breast Cancer Conditional Outcome Calculator
CancerMath
Breast Cancer Tools
All Cancers
About
Outcome
Conditional Survival
Therapy
Nodal
Nipple Involvement
Enter patient information:
Factors affecting non-cancer lethality
Current Age:
61
Factors affecting cancer lethality
Years Since Diagnosis:
1
Evidence of recurrence:
Unknown
No
Tumor Diameter (cm):
1
# of Positive Nodes:
Known
1
ER Status:
Positive
PR Status:
Positive
HER2 Status:
Positive
Histological Type:
Ductal
Grade:
1 - Well diff.
Update Graph
Questions or trouble? Click here for the calculator FAQ
mortality risk
enlarge
Cancer
Non-cancer
Overall
Year after diagnosis
Display as:
Mortality curves
Cancer Mortality:
9.5% expected remaining cancer death rate.
Life Expectancy:
The remaining chance of cancer death cancer shortens the life expectancy of a 61-year-old woman by 1.4 years. (from 23.1 years to 21.7 years)
Laboratory for Quantitative Medicine
LifeMath.net
Copyright © 2007-11 James Michaelson, PhD
Legal Disclaimer (read before using site)

Chapter 2 Aims and objectives of the thesis

The overall aim of the thesis is to determine long term survival among different subgroups of breast cancer patients, to evaluate various risk factors on outcome as well as to assess the performance of prognostic models in Southeast Asia.

2.1 Study 1 – Incidence and outcome of male breast cancer: an international population-based study (Chapter 5)

Increasing incidence rate and improved survival of female breast cancer have been reported in many studies. Due to the low incidence, good recent data on risk and outcome of male breast cancer is lacking and most studies suffered from small sample sizes, short follow-up and non–population-based designs. This study aims to improve our understanding of risk and outcome of male breast cancer in relation to female breast cancer, to:

i) Compare the period trend in incidence between both genders

ii) Compare the period trend in survival between both genders

iii) Compare the relative survival between both genders

2.2. Study 2 – The impact of *in situ* breast cancer and family history on risk of subsequent breast cancer events and mortality: a population-based study from Sweden (Chapter 6)

There is lack of evidence on long term outcome of breast carcinoma *in situ* and the aims of this study are to:

i) Assess the risk of subsequent invasive breast cancer and mortality in women diagnosed with *in situ* breast cancer

ii) Assess the change of risk of subsequent invasive breast cancer over time which might be attributed to advancement in detection and treatment.

iii) Evaluate the effect of family history, age at diagnosis, follow-up time on risk of subsequent invasive breast cancer and mortality after diagnosis of breast carcinoma *in situ*

2.3 Study 3 – Validation of the CancerMath prognostic tool for breast cancer in Southeast Asia (Chapter 7)

The most commonly used prognostic prediction programme Adjuvant! Online overestimates survival in Asian patients. We aim to validate another prognostic model CancerMath for early breast cancer patients in Southeast Asia.

2.4 Study 4 – Predicting survival of *de novo* metastatic breast cancer in Asian women: systematic review and validation study of prognostic tools. (Chapter 8)

In Asia, up to 25% of breast cancer patients present with distant metastases at diagnosis. Given the heterogeneous survival probabilities of *de novo* metastatic breast cancer, individual outcome prediction is challenging. The aim of the study is to identify existing prognostic models for patients with *de novo* metastatic breast cancer and validate them in Asia.

Chapter 3 Source of data and study design

In epidemiological studies, it is important to define the study population and study design in advance as it will directly affect the validity and generalizability of the findings. In this chapter, we discuss the study design and source of data in detail.

3.1 Overview of epidemiological study designs

Epidemiological studies can be either experimental or observational, depending on whether the researchers assign exposure to the study subjects or simply observe the effect of exposure. For etiological research, the aim is to investigate a causal relation between a risk factor and a health-related event. Confounding is a central issue when establishing causality because the true association can be distorted by another variable. One method to control for confounding is through random allocation of exposure in an experimental study so that the comparison groups are as similar as possible. Such study design is known as randomized controlled trial (RCT). RCTs and pooled analysis of RCTs provide the highest quality of evidence as it eliminates both known and unknown confounding bias [221]. However in this thesis, RCT was not feasible in study 1 and 2 as certain risk factors cannot be randomly administered to patients, such as gender, time of diagnosis and family history. For prognostic research studies such as study 3 and 4, the aim was to predict the outcome in a multivariate manner instead of establishing the causal relation. In this type of studies, confounding is not an issue. RCTs are not appropriate for prognostic studies, except for impact studies which aim to measure the effect of using a prognostic model on doctors' behavior, patient outcome, or cost effectiveness of care compared with not using such model [222-224].

Well-designed observational studies, mainly cohort study and case-control study can also provide high level of evidence [221]. For prognostic research where RCT is not appropriate, prospective cohort study provide the highest level of evidence [225]. In a cohort study, a group of people who are at risk of a specific outcome are selected based on their exposure status and followed over a period of time to determine outcome. The cohort can be followed prospectively for events to happen or the experience of a cohort over a period of time can be reconstructed retrospectively using historical records when outcome has already occurred. In both prospective and retrospective cohorts, exposure status is determined before the occurrence of outcome. Therefore a temporal relationship can be potentially established to demonstrate causal association. Prospective data collection can be very time consuming and costly to have sufficient number of events. In retrospective studies, information on exposure and potential confounders in existing records might be incomplete and inaccurate[226]. Other disadvantages of cohort studies include loss of follow-up and differential misclassification of outcome between exposed and non-exposed groups. For example, women who take oral contraceptives are monitored and examined more carefully for conditions like hypertension and venous thrombosis thus have higher probability of detection than non-users [227, 228].

Another common observational study design is case-control study, in which the frequency of exposure is compared between individuals with (cases) and without (controls) a particular outcome. To minimize selection bias, the cases and controls should come from the same source population. Cases and controls can be matched based on certain variables to account for confounders. In contrast to cohort study, the outcome status is already known and exposure status is determined after

occurrence of outcome, which can introduce recall bias as cases are more likely to report exposure[226]. Although case-control study does not allow estimation of incidence and is more susceptible to recall bias and interview bias, it is quick, inexpensive and more efficient for rare outcomes [229]. The last common study design of observational study is cross sectional study, which is conducted at one time point to usually estimate the prevalence of certain outcome[230]. Therefore it cannot assess the causal effect and cannot be used for outcome and prognostic research.

In this thesis, all four studies are retrospective cohort studies by using existing population-based or hospital-based database.

3.2 Consolidation of population-based cancer registry data – Study 1

Six population-based cancer registries in Denmark, Finland, Geneva (Switzerland), Norway, Singapore, and Sweden contributed incident cases to the study on male breast cancer.

The Danish Cancer Registry

The Danish Cancer Registry was founded in May 1942 and started registering all incidences of malignant neoplasms and certain benign lesions from Danish residents in the following year [231]. Reporting to the cancer registry has become mandatory since 1987. For unknown cancers not notified by hospital departments, pathology departments or physicians in general practice and only identified in death certificates, detailed information can be obtained from the national death certificate system [232]. Such cases have reduced from 19% in the 1940s to 1-2% in the 1980s. The completeness and validity of the registry was assessed to be 95% to 98% in

1977 [233]. The proportion of morphologically verified cases was 89% [231]. TNM stage information is available only after 2004, and the completeness of TNM registration for breast cancer was 85.4% [234]. Information on emigration, immigration and death was updated once a year by linkage with the Danish Civil Registration System and the Danish Register of Causes of Death using the ten-digit unique Civil Personal Register number (personal identity number) allocated for each resident at birth or when obtaining permanent residence [231].

The Finnish Cancer Registry

The Finnish Cancer Registry was founded in 1952 and all hospitals, physicians and pathological and haematological laboratories must send a notification of every cancer case from 1961 onwards [233]. The completeness of registration of solid tumors was reported to be over 99% [235]. In 2003-2007, 93% of cancer cases were register based on microscopic verification and 2% were based only on death certificates [236]. Data on stage and basic treatment has been recorded since the beginning of cancer registration. Vital status, date and causes of death are annually matched with the record from Cause of Death Register and Central Population Register.

The Geneva Cancer Registry

The registration of cancer in Switzerland is managed by cantonal level registers [237]. The Geneva Cancer Registry records all incident cancers occurring in the population of the canton of Geneva since 1970[238]. All hospitals, pathology laboratories, and private practitioners are required to notify all cancer cases [238]. Completeness of registration was reported to be very high with less than 2% of cases were recorded from death certificates only [239]. Information on stage at

diagnosis, hormone receptor status and treatment within 6 months after diagnosis is included in the registry [238]. Cause of death is obtained from death certificate or hospital records as well as annual active follow-up via linkage with Cantonal Population Office [240].

The Norwegian Cancer Registry

The Cancer Registry of Norway has collected cancer notifications since 1952. All hospitals, laboratories and general practitioners are legally obligated to report new precancerous and cancerous cases within two months of diagnosis [241, 242]. Estimated completeness of the registry was 98.8% for the period of 2001 to 2005 and 93.8% of cases were morphologically verified [242]. The incidence registry includes basic data collected from clinical and pathological reports, patients' discharge and mortality records [241]. Clinical registries with detailed information on diagnosis, pathological examination, treatment and follow up were also established for several major types of cancer [241]. Death records come from National Population Registry and are compared with data from the Cause of Death Register run by Statistics Norway at least once a year.

The Singapore Cancer Registry

The Singapore Cancer Registry started registration of incident cancers in January 1968 and notification became compulsory in 2009[174]. Multiple sources of notifications such as medical profession, pathological reports, discharge records and death certificates are used to ensure the registry is as complete as possible. The proportion of cases by death certificate only was 4.2% between 1968 and 1997, 1.0% for the period 1993-1997 and 0.9% for the period 1998-2002 [243, 244]. Vital status and cause of death is retrieved through linkage to the death registry and nearly

all deaths are certified in Singapore [174]. Information on stage is available from 2003 onwards, whereas no information on treatment is available.

The Swedish Cancer Registry

The Swedish Cancer Registry has registered newly detected cancer cases since 1958 and registration, coding and quality check has been performed at six regional registries since the mid-1980s[245]. It is compulsory for all healthcare providers to notify new cancer case and approximately 99% of the cases are morphologically verified. Less than 2% of cases were not reported in the late 1970s and recent study showed underreporting of 3.7% of cases in 1998 based on records from Hospital Discharge Register [245, 246]. Cases reported by death certificates only are not included in the registry[233]. Vital status and date of death and migration is obtained via linkages to the Cause of Death Register and the Total Population Register. Stage has been collected since 2004 but was not included in our analysis. There is no information on treatment.

3.3 Linkage of multiple population-based registries in Sweden – Study 2

The Multi-Generation Register (MGR) in Sweden includes more than 10 million Swedish residents who are born after 1932 in Sweden and have ever been registered in Sweden after 1st January 1961[247]. These people are called index persons. Basic information on index person such as date of birth and country of birth and on their parents is taken from three registers in Sweden, i.e. the Total Population Register, Personal Records and Statistics Sweden's register of births. The coverage is comprehensive for index person except for missing data on those emigrated between 1961 and 1967 and never returned. Information on both biological and

adoptive parents of index persons is recorded and 98% maternal information and 95% paternal information is complete for the Sweden-born subpopulation [247, 248]. This register is updated in March every year and enables identification of parents, siblings, children and other relatives of the index person. Via the unique national registration number given at birth or immigration, information of index person and his/her relatives can be retrieved from other registers at Statistics Sweden or other authorities in Sweden.

The Swedish Cancer Registry recorded DCIS as invasive cancer until 1980 and since then both DCIS and LCIS were classified as *in situ* [249]. The register does not distinguish ductal from lobular *in situ* breast cancer before 1990 and has no information regarding tumor stage or treatment. The completeness of registration of *in situ* cases improved from 78% in the 1980s to 95% in early 1990s and the correctness improved from 94% to 96% [250]. For the second study, we linked the data from the MGR with the Swedish Cancer Registry using the unique national registration number and identified women in MGR diagnosed with breast carcinoma *in situ* to be included in the analysis.

3.4 The Singapore Malaysia Hospital Based Breast Cancer Registry – Study 3 and Study 4

Singapore Malaysia Hospital Based Breast Cancer Registry consists of three hospital-based breast cancer registries in Singapore and Malaysia. National University Hospital (NUH) and Tan Tock Seng Hospital (TTSH) are two public tertiary hospitals in Singapore. The breast cancer registry at National University Hospital (NUH) in Singapore contributed cases diagnosed since 1990 whereas the Tan Tock Seng Hospital (TTSH) registry included patients diagnosed from 2001

onwards. University Malaya Medical Centre (UMMC), located in Kuala Lumpur, Malaysia, has prospectively collected data on breast cancer cases since 1993. These registries have received approval from respective ethical review committees. All three registries include data on basic patient demography, age and date of diagnosis, histologically determined tumor size, number of positive lymph nodes, ER and PR status (positive defined as 1% or more positively stained tumor cells at NUH or 10% or more positively stained tumor cells at UMMC and TTSH, negative, or unknown), HER2 status based on FISH and IHC if FISH is not performed (positive defined as FISH positive or IHC score of 3+, negative defined as FISH negative or IHC scored of 0 or 1+, equivocal defined as IHC score of 2+, or unknown), histological type (ductal, lobular, mucinous, others, or unknown), grade (1, 2, 3, or unknown), type of surgery (no surgery, mastectomy, breast conserving surgery, or unknown), chemotherapy (yes, no or unknown), hormone therapy (yes, no, or unknown), and radiotherapy (yes, no, or unknown). Stage I to IV was categorized according to the sixth edition of AJCC staging system. *De novo* metastasis was defined as distant metastasis detected within three months after diagnosis and metastasis in the ipsilateral supraclavicular lymph nodes only was not considered as metastatic patients according to the sixth edition. Site(s) of metastasis (bone, lung, liver, brain, soft tissue or other organ) was recorded for *de novo* metastatic breast cancer patients at all three registries. Detailed chemotherapeutic treatment regimens were only available for UMMC patients. For chemotherapy, cyclophosphamide methotrexate fluorouracil (CMF) was categorized as first generation regimen and fluorouracil, epirubicin and cyclophosphamide (FEC), and doxorubicin and cyclophosphamide (AC) followed by paclitaxel were second generation. Docetaxel, doxorubicin and cyclophosphamide (TAC), and FEC followed by doxorubicin were categorized as

third generation. Hormone therapy was categorized into five groups: tamoxifen, aromatase inhibitors (AI), tamoxifen followed by AI, ovarian ablation, and ovarian ablation plus tamoxifen. Death information was obtained from the hospitals' medical records and ascertained by linkage to death registries in both countries.

For study 3, Women diagnosed with pathological stage I to III breast cancer between 1990 and 2011 were identified from the registry. For study 4, *de novo* metastatic breast cancer patients diagnosed between 2000 and 2010 were identified from this registry.

3.5 Systematic review – Study 4

Systematic review is a strategy to gather best available evidence on a specific research question and the results from each study on this topic can be pooled and analyzed by an explicit method known as meta-analysis (not performed in this thesis). In the fourth study in this thesis, a systematic review on prognostic tools to predict overall survival of metastatic breast cancer patients was conducted.

Before initiation of a systematic search of relevant studies, the research question must be clearly structured, defining the study population, outcome of interest, exposure/intervention, and study design without ambiguity [251]. To ensure the quality of systematic review, PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) guidelines were developed by an international group of multidisciplinary experts [252]. Inclusion and exclusion of studies should follow the study selection criteria specified earlier and reasons for exclusion should be recorded. Evidence from selected studies can be summarized qualitatively and

quantitatively (not done in this thesis). It is also important to assess the quality of the studies as it may potentially affect the interpretation of the results.

Recently a more specific checklist for systematic reviews of prediction modelling studies was designed to guide formulation of research question and appraisal of prediction modelling studies [253]. The type of prediction modelling studies for which this checklist can be applied to includes development with or without external validation and external validation studies with or without modification of the original model. This checklist can be used for both diagnostic and prognostic prediction model and it covers 11 domains, including source of data, participants, outcomes, candidate predictors, sample size, missing data, model development, performance, evaluation, results, and discussion. The systematic review conducted in the fourth study has extracted relevant items listed in each domain from studies included in the review.

Chapter 4 Statistical analysis

This chapter gives a brief introduction of the statistical techniques used and discussed in this thesis. The detailed analytical plan for each study can be found in Chapter 5-8.

4.1 Survival analysis

To study the outcome of breast cancer, we have to follow the patient from diagnosis of breast cancer to occurrence of a particular event, such as death or recurrence of cancer. The duration from time of diagnosis to the event is called the "time-to-event" or "survival time". In some circumstances such as the event of interest has not occurred when the study ends or when the patient dies, or event status is not known due to lost-to-follow-up, the survival time is right censored. We then use the observed survival time to draw implication about the true survival time [254].

In most studies, we are interested to know the probability of survival after breast cancer diagnosis. The simplest way is to present the proportion of subjects whose survival time exceeds a fix period of time (e.g. 5-year). However it does not take observations that are censored during the first 5 years into consideration and does not fully utilize the exact survival time. Two methods were developed to deal with censoring data: life table method and Kaplan-Meier method. Both methods assume that subjects who are censored have the same survival experience as those who are followed.

Life table (actuarial) method

The proportion of dying (q_i) and proportion of survival (p_i) are calculated for each fixed time interval such as every one year after diagnosed by q_i= number of patients dying during that time interval/effective number exposed to risk of dying and p_i=1-q_i [255]. The effective number at risk of dying (r_i) takes both number of patients alive at beginning of that interval (l_i) and number of patients last seen alive (censored) during that interval (w_i) in to account by assuming censored patients on average were followed up for half of the interval, which results in $r_i=l_i-w_i/2$. The cumulative survival rate until end of n^{th} interval is then the product of p_i from 1^{st} to n^{th} interval.

Kaplan-Meier method

The Kaplan-Meier method to estimate survival probability is very similar to the life table approach. The only difference is that the probability can be calculated whenever a death occurs instead of at the end of a fixed time interval. As a result, the interval length is no longer consistent as comparing to the life table approach. The conditional probability of surviving the time interval (can be measured in days, months or years depends on how precise the data is) given being alive at beginning of the interval is calculated as number of patients survived by the end of internal divided by number of patients alive at beginning of that interval. Censored cases are considered to have survived throughout the time interval. The surviving probability up to certain time point is then the product of conditional probabilities of previous intervals (product limit method). Standard error of the estimates can be computed by $\sqrt{\frac{P(1-P)}{N}}$ where P is the cumulative survival and N is number of subjects.

The Kaplan-Meier survival curve can be plotted based on estimate at end of each interval. The survival curve starts at value of 1 (or 100%) and proceeds horizontally until an event occurs. The depth of drop depends on how many events occur at that point of time. Two or more Kaplan-Meier survival curves of different subgroups such as patients with and without treatment can be compared using log-rank test, which is a chi-square test with degree of freedom equals to number of groups-1. The null hypothesis is that there is no difference between the groups at any time point. At each time (j) an event occurs, observed (Oij) and expected (Eij) number of events are calculated for each group i. The test statistics can be approximated by $\sum\frac{(Oi-Ei)^2}{Ei}$ where and Oi=∑Oij and Ei=∑Eij.

Relative survival

Overall survival estimates the probability surviving all causes of death. Sometimes we are only interested in net probability of survival with cancer as the only cause of death. Net survival (or excess mortality) can be estimated by cause-specific survival where death from other causes other than cancer of interest is censored. However cause of death is difficult to determine when metastasis or treatment complication occurs and death certificates are often unreliable or unavailable for population-based study [256]. Relative survival is used as another measure of net survival by calculating the ratio of observed survival of cancer patients to the expected survival of a comparable cohort from the general population, usually matched by gender, age and calendar period [257]. Observed survival of the cancer patients is estimated using the life table approach and expected survival can be derived from annual probability of death reported in population life tables (stratified by gender- and calendar period) according to Ederer I [257], Ederer II

[258] or Hakulinen method [259]. Population life table or annual probability of death is usually published and publically available as part of national statistics in many countries. They can also be downloaded from the Human Life-table Database (http://www.lifetable.de/) and the Human Mortality Database (http://www.mortality.org/). The expected survival estimated from these three methods does not differ much except for long term survival of cancer sites which affect people from wide range of age groups. Ederer II method is preferred in many cancer registries as the relative survival calculated is lower than the other two methods.

Standardized incidence/mortality ratio

Occurrence of cancer relapse or death in a particular population can be compared with a reference population using standardized incidence/mortality ratio (SIR and SMR). The ratio is calculated as observed number of events (recurrence if SIR or death if SMR) divided by the expected number of events, where expected number is incidence/mortality rate of the reference population multiplied by the person-time of the population of interest. Expected number can be estimated for each gender-, age-, period-specific group and then added up therefore the ratio is standardized to adjust for different age distributions of the two populations. An SIR/SMR of 1 indicates no difference of incidence/mortality between the two populations. Confidence interval (CI) is calculated for SIR/SMR by assuming observed number of events is Poisson distributed to determine whether the difference is significant [260, 261].

4.2 Regression analysis

The relationship between one dependent (response) variable and a group of independent (explanatory) variables can be modelled statistically using regression analysis. The simplest regression model is linear regression, where the relationship between dependent variable Y and independent variables X_i is assumed to be linear, and can expressed as $Y=\beta_0 + \beta_1X_1 + \beta_2X_2+ \ldots+\beta_nX_n +\varepsilon$. The coefficient β_i shows how much Y will change when X_i increase by 1 unit.

In epidemiological studies, the effect size of risk factor on outcome can be estimated using regression model while adjusting for other confounders. In prognostic research, several predictors can be combined mathematically using regression model to predict outcome. In this section, we will discuss several regression methods used and/or discussed in this thesis.

Cox proportional hazard regression model

In previous section, we have demonstrated survival time of two or more groups can be compared using Kaplan-Meier curves and log-rank test. However this analysis is univariate and is only suitable for categorical variables. Cox regression is developed to study the joint effect of multiple covariates and to estimate the effect of one risk factor while adjusting for confounders. The dependent variable Y in a Cox regression model is the hazard function at a given time t, denoted $\lambda(t)$, which can be considered as the instantaneous rate that an event occurs at time t [262, 263]. It can be related to various explanatory variables as $\lambda(t) = \lambda_0(t)\exp(\beta_1X_1 + \ldots+\beta_nX_n)$, and $\lambda_0(t)$ is the baseline hazard function when all the explanatory variables are zero. The ratio of two hazard functions is known as hazard ratio and is assumed to be constant over time, i.e. the hazard of one group is proportional to the other group

over time. This assumption can be tested using complementary log-log plot. Estimates of β_i and CI can be obtained by method of partial likelihood. The hazard ratio comparing $X_i=x_i$ to $X_i=x_{i'}$ is then $\exp(\beta_i(x_i- x_{i'}))$. The hypothesis of equal hazard，i.e. $\exp(\beta_i)=1$, or $\beta_i =0$, can be tested using partial likelihood ratio test. In prognostic research, weighted sum of the covariates in the Cox model, where the weights are the coefficients β_is, is always used as prognostic index/score [264].

Poisson regression

Although Cox regression is widely used in survival analysis, it does not accommodate multiple time scales simultaneously (e.g. attained age, time since diagnosis, and calendar time) and cannot be applied to model a difference in two rates [265, 266]. Poisson regression is used as an alternative method. It is called Poisson regression because number of events occurred (Y) is modelled by generalized linear model with log link function under the assumption of Poisson distribution. The logarithm of expected value of Y, denoted $\log(\mu)$ can be modelled as $\log(\mu) = \beta_0 + \beta_1X_1 + \ldots+\beta_nX_n$. The dependent variable can also be incidence/mortality rate Y/t (t is the accumulated person-time at risk) and the equation becomes $\log(\mu/t) = \beta_0 + \beta_1X_1 + \ldots+\beta_nX_n$, . It can be further modified as $\log(\mu)= \log(t)+ \beta_0 + \beta_1X_1 +\ldots+\beta_nX_n$, where log(t) is known as the offset. Based on the model, we can calculate $\exp(\beta_i(x_i- x_{i'}))$ as the incidence/mortality rate ratio comparing $X_i=x_i$ to $X_i=x_{i'}$. All the covariates has multiplicative effect on incidence/mortality rate as $\mu/t= \exp(\beta_0)\exp(\beta_1X_1).. \exp(\beta_nX_n)$.

Relative excess risk

As discussed earlier, relative survival is a measure of net survival (or excess mortality) attributable to cancer. To evaluate effect of one risk factor on relative survival while controlling for other confounders, the excess mortality (observed mortality minus expected mortality) can be modelled under the assumption of Poisson distribution as a multiplicative function of covariates and offset by logarithm of person-time at risk, $\log(\mu - d*) = \ln(t) + \beta_1 X_1 + \ldots + \beta_n X_n m$, where $d*$ is expected number of death. The ratio of excess mortality comparing $X_i = x_i$ to $X_i = x_{i'}$ (relative excess risk, RER) is then $\exp(\beta_i(x_i - x_{i'}))$, which is the same as rate ratio in Poisson regression.

Additive Poisson model and excess additive risk

In both Cox regression and previous examples of Poisson regression, effects of different factors on risk are combined multiplicatively, i.e. risk ratio of having multiple risk factors is a production of individual risk ratio for each risk factor. In some situation, an additive effect, where the risk differences from different factors are added together, is more appropriate. For count data with person time, rate can be modelled using identify link instead of natural log link, i.e. $\mu/t = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n$.

Using this method, the absolute difference of observed and expected mortality rate, known as excess additive risk (EAR) can be estimated using a Poisson additive model with expected number of cases as the offset. A likelihood ratio test was used to calculate 95% CI.

4.3 Validation of prognostic model

The third and fourth studies in this thesis validated various prognostic models developed for breast cancer patients. The predicted outcome from these prognostic models is usually an absolute risk of an event or a prognostic score [267]. Most of prognostic models are developed using a limited sample size and having many potential predictors being tested. The resulting model may be over-fitted to the derivation data and shows optimistic performance when it is applied to data from the same source as the development set [268, 269]. Although internal validation using split-sample, cross validation or bootstrapping can correct the optimism of model performance [270], it is limited in terms of assessing generalizability of the model [269]. Therefore it is important to validate prognostic models in an external dataset, either from different time period or different geographical area. The predictive performance of the prognostic model can be evaluated in terms of its discrimination and calibration.

Discrimination

Discriminative ability of a prognostic model describes how well the model distinguishes between patients with good and poor outcome. If the outcome of interest is dichotomous (with or without event of interest), different cut-off levels can be applied to the prognostic score/predicted probability to classify a patient as positive or negative for outcome. For each possible threshold, sensitivity and specificity can be calculated. Each pair of sensitivity and 1-specificity can be plotted against each other and the resulting curve is known as the receiver operating characteristic (ROC) curve. The area under a ROC curve (AUC) is interpreted as the probability of assigning a higher prognostic score or predicted risk to a

randomly selected individual with the outcome of interest than to another randomly selected individual without the outcome. For survival models, length of follow up should be specified so that dichotomous survival status can be used for ROC analysis [271]. However censored cases cannot be included in analysis due to their unknown status. Another adapted method was developed by Harrel to compare the survival time between any possible random pair of subjects [272]. The probability of assigning greater prognostic score to a person who survived longer is called concordance statistic (C-statistic). Same as AUC, a C-statistic of 0.5 indicates no discrimination and value of 1.0 means perfect discrimination.

Discrimination can be affected by the heterogeneity of the validation population. As the spread of the predicted probability or score increases in the validation set, the model tends to discriminate better [273]. Therefore the difference in underlying risk distribution between development and validation sets could affect its discrimination in the validation study.

Calibration

Calibration refers to the agreement between the predicted outcome and observed outcome, which reflects the accuracy of the prediction. It can be assessed by splitting the data into several groups, normally based on deciles of predicted outcome, and then comparing the predicted outcome and observed outcome in each decile. For example, for binary outcome, the proportion of events can be compared with average or median predicted probability for each group. A calibration plot is presented by plotting the observed outcome against predicted outcome for each group. The 45 degree diagonal line illustrates perfect agreement between predicted and observed outcome. The observed number of events can also be compared with

predicted number of events in each group (sum of predicted probability) using Hosmer–Lemeshow test. In contrast to discrimination, calibration is affected by the differences in means of predicted outcome between development and validation sets [273]. If persistent under-prediction or over-prediction is observed across all groups, it is necessary to recalibrate the model.

Chapter 5 Incidence and outcome of male breast cancer: an international population-based study

Originally published by the American Society of Clinical Oncology. [Miao H, Verkooijen HM, Chia KS, Bouchardy C, Pukkala E, Larønningen S, Mellemkjær L, Czene K, Hartman M: Journal of Clinical Oncology 29 (33), 2011 Nov 20: 4381-6]

5.1 Motivation

Male breast cancer is a rare disease, accounts for 0.5% to 1% of all breast cancer cases [11, 274-276]. Similar to female breast cancer, the risk of male breast cancer increases steadily with age [11, 277], although men are, on average diagnosed at later ages and do not display the typical deceleration in risk after the age of 50 years as seen in women, described by Clemmesen over 60 years ago [11, 13, 274, 277-279].

Male breast cancer is reportedly associated with worse outcome as compared to female breast cancer [126, 276]. Some studies have suggested that survival differences between genders disappear after stratification for age and stage [126, 278, 280]. However, given the low incidence of male breast cancer, many of these studies suffered from small sample sizes, short follow-up time and a non-population-based design, limiting their interpretability.

Over the last few decades, survival of female breast cancer has improved substantially. This is likely a combined result of earlier detection and improvements in treatment [7, 281]. Given the scarcity of male breast cancer, solid recent data on risk and outcome for male disease is lacking. We have undertaken a population-

based international study, with the aim to improve our understanding of risk and outcome of male breast cancer in relation to female breast cancer.

5.2 Methods

For the current study, we included patients with invasive breast cancer from all six participating regions diagnosed between 1970 and 2007 with the exception of Denmark, where patients diagnosed up to 2006 were included. Detailed description of cancer registries in these six regions can be found in Chapter 3. All datasets contained information on sex, date of birth, date of diagnosis, duration of follow-up, vital status, date of death and date of migration for all individuals. Stage at diagnosis was available for patients diagnosed in Finland (localized, regional, distant, or unknown). For patients from Geneva, Norway and Singapore, TNM stage was transformed to localized, regional, or distant, with stage I as localized, stage II and III as regional, and stage IV as distant. Basic treatment information was available for Finland, Geneva and Norway and included surgery (yes, no, or unknown), chemotherapy (yes, no, or unknown), radiotherapy (yes, no, or unknown) and hormonal therapy (yes, no, or unknown).

Patients with an invasive cancer diagnosis before first breast cancer were excluded, as were individuals who immigrated from another region before diagnosis, because of the possibility of misclassification of cancer history. For individuals with multiple breast cancer diagnoses, only the first cancer was included in analysis. Our final study population comprised 459,846 women and 2,665 men diagnosed with invasive breast cancer. Follow-up started at time of diagnosis, and survival time was defined as the time between the date of diagnosis and date of death, emigration, or end of follow-up (December 31, 2007), whichever occurred first.

Statistical analysis

Mann-Whitney U test and Chi-square test were performed to test gender differences in distribution of age, stage and treatment. Significance level of the associations was based on valid proportions only (ie, after excluding missing information). The age standardized incidence rate of invasive breast cancer was calculated using the total female and male population of the six regions as denominators and was directly standardized to world standard population with 5-year age groups.

Overall survival was evaluated using Kaplan-Meier analysis stratified by gender and stage. We applied relative survival analysis to account for differences in life expectancy between men and women. Overall relative survival ratios (RSRs) for both genders were estimated at 5 and 15 years follow-up. To investigate improvements in relative survival over calendar time for men and women, we evaluated trends in 5-year RSR by stage and over time (10-year categories).

To adjust for potential confounders such as age at diagnosis, calendar period of diagnosis (grouped by every five years), follow-up time (group by every one year), region, stage and treatment, we modelled the excess risk using Poisson regression. The reference category for gender comparison was the female group. Regression models were built using three datasets: all individuals (i.e., analysis including individuals from all regions), individuals from regions with stage information (i.e. Finland, Geneva, Norway and Singapore) and individuals from regions with both stage and treatment information (i.e. Finland, Geneva and Norway). On the basis of the latter dataset, we stepwise evaluated the effect of adjustment for age, stage, and treatment on the relative risk of death from breast cancer for males compared to females. Statistical analysis was done using the SAS Statistical package, version 9.2.

5.3 Results

Male breast cancer (n=2,665) represented 0.6% of all breast cancers (Table 5.1), and this proportion was similar for all 6 regions. Women were diagnosed with breast cancer at a younger median age than men (61.7 vs. 69.6 years, respectively; p<0.001). Among the 190,030 (41%) breast cancer patients with information on stage, 41% of the men and 44% of women were classified as having localized disease. Distant disease extent accounted for 11% and 6% for men and women respectively (p<0.001). For 167,169 patients (36%) with information on treatment, men were significantly less likely to receive surgery and radiotherapy, but there were no differences in the administration of chemotherapy and hormonal therapy.

Table 5.1 Characteristics of female and male breast cancer cases diagnosed in Denmark, Finland, Geneva, Norway, Singapore and Sweden between 1970 and 2007 ††

Characteristics	***Male N= 2,665, 0.6%***	***Female N=459,846, 99.4%***	***P-value***
Region			<.001*
Denmark	677 (25.4%)	97,228 (21.1%)	
Finland	347 (13.0%)	86,083 (18.7%)	
Geneva	61 (2.3%)	9,980 (2.2%)	
Norway	435 (16.3%)	70,263 (15.3%)	
Singapore	74 (2.8%)	22,787 (5.0%)	
Sweden	1,071 (40.2%)	173,505 (37.7%)	
Age, years			
Median age	69.6	61.7	<.001¶
0-40	62 (2.3%)	25,154 (5.5%)	<.001*
40-60	612 (23.0%)	185,901 (40.4%)	
60+	1,991 (74.7%)	248,791 (54.1%)	
Calendar Period			<.001*
1970-1977	490 (18.4%)	67,478 (14.7%)	
1978-1987	607 (22.8%)	101,755 (22.1%)	
1988-1997	728 (27.3%)	130,029 (28.3%)	
1998-2007	840 (31.5%)	160,584 (35.0%)	
Stage			<.001†
Localized	379 (41.3%)	83,828 (44.3%)	
Regional	311 (33.9%)	64,945 (34.3%)	

Distant	100 (10.9%)	10,561 (5.6%)	
Unknown	127 (13.9%)	29,779 (15.8%)	
Total	917	189,113	
Treatment			
Surgery			<.001‡
Yes	728 (86.4%)	150,769 (90.7%)	
No	79 (9.4%)	9,572 (5.8%)	
Unknown	36 (4.3%)	5,985 (3.6%)	
Radiotherapy			<.001‡
Yes	251 (29.8%)	63,751 (38.3%)	
No	447 (53.0%)	81,775 (49.2%)	
Unknown	145 (17.2%)	20,800 (12.5%)	
Chemotherapy			0.06‡
Yes	127 (15.1%)	31,125 (18.7%)	
No	542 (64.3%)	110,749 (66.6%)	
Unknown	174 (20.6%)	24,452 (14.7%)	
Hormonal therapy			0.09‡
Yes	190 (22.5%)	35,400 (21.3%)	
No	508 (60.3%)	109,199 (65.7%)	
Unknown	145 (17.2%)	21,727 (13.1%)	
Total	843 (100%)	166,326 (100%)	

†† Denmark contributed case diagnosed between 1970 and 2006 * Chi-square test ¶ Mann-Whitney U test † Chi-square test on valid proportion only, using subset with stage information (Finland, Geneva, Norway and Singapore) ‡ Chi-square test on valid proportion only, using subset with treatment information (Finland, Geneva and Norway)

The overall age standardized incidence rates were 0.4 per 100,000 person-years in men and 66.7 per 100,000 person-years in women. The incidence of breast cancer in women increased by more than 50%, from 51.4 per 100,000 person-years in the early 1970s to 80.3 per 100,000 person-years after the year 2000 (Figure 5.1). The overall incidence of disease in men remained stable at approximately 0.4 per 100,000 person-years all the time (Figure 5.1). Compared with the European countries, Singapore had a lower incidence rate for both genders, but a faster increase in incidence, which tripled in women and quadrupled in men from the early 1970s to the 2000s (female: 23.97 to 63.17 per 100,000 person-years; male: 0.05 to

0.21 per 100,000 person years) (Figure 5.2 and 5.3). The increased incidence for men in Singapore was not statistically significant. Men had a worse overall survival compared with women (Figure 5.4), except for patients with distant spread of disease, for whom overall survival was similar for both genders. Disease-specific survival (as estimated by relative survival) was significantly worse for male patients at both 5 and 15 years compared with female patients (5-year RSR, 0.72 vs 0.78 respectively; 15-year RSR 0.50 vs. 0.61 respectively) (Table 5.2). This corresponds to a 27% higher unadjusted 5-year excess mortality risk for men compared with women (RER=1.27, 95% CI, 1.13-1.42; Table 5.2, model 1a) and a 36% higher 15-year excess risk (RER=1.36, 95% CI, 1.24-1.50; Table 5.2, model 1a). After adjusting for region, age and year of diagnosis, follow-up time, and stage, there was no significant difference in 5-year and 15-year excess mortality between men and women (Table 5.2, model 3a and 3b). Additional adjustment for treatment further reduced the RER to 0.78 (95% CI, 0.62-0.97; Table 5.2, model 4). A similar pattern was observed when assessing 15-year follow-up. For female patients with breast cancer, 5-year relative survival increased from 0.66 (95% CI, 0.66-0.67) in 1970 to 1977 to 0.87 (95% CI, 0.86-0.87) in 1998 to 2007 (Table 5.3). Men experienced an improvement in relative survival as well, from 0.67 (95% CI, 0.60-0.72) in 1970 to 1977 to 0.78 (95% CI, 0.73-0.83) in 1998 to 2007.

Figure 5.1 Incidence rate (IR) of male and female invasive breast cancer (standardized to world population), by period of diagnosis (right y-axis denotes males, left y-axis denotes females)

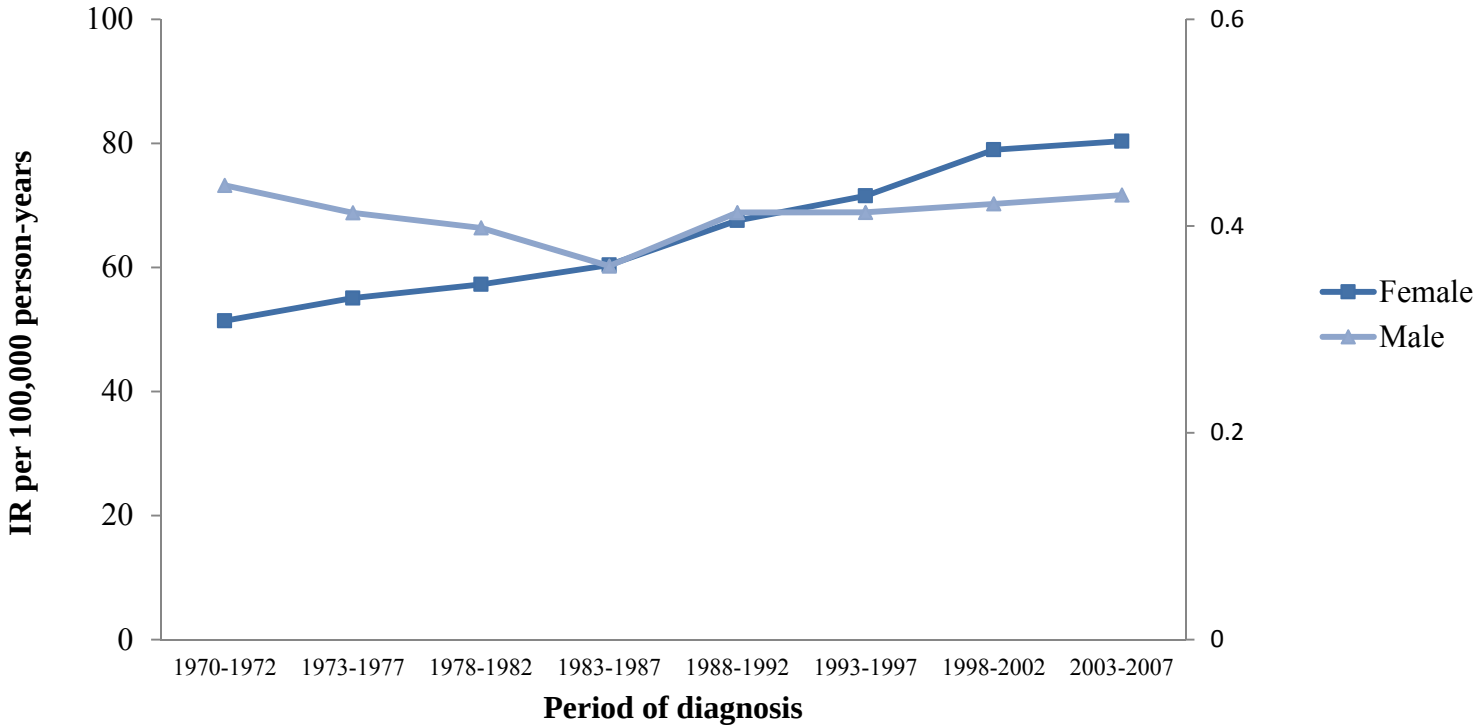


Figure 5.2 Incidence rate (IR) of male invasive breast cancer (standardized to world population), by period of diagnosis and region

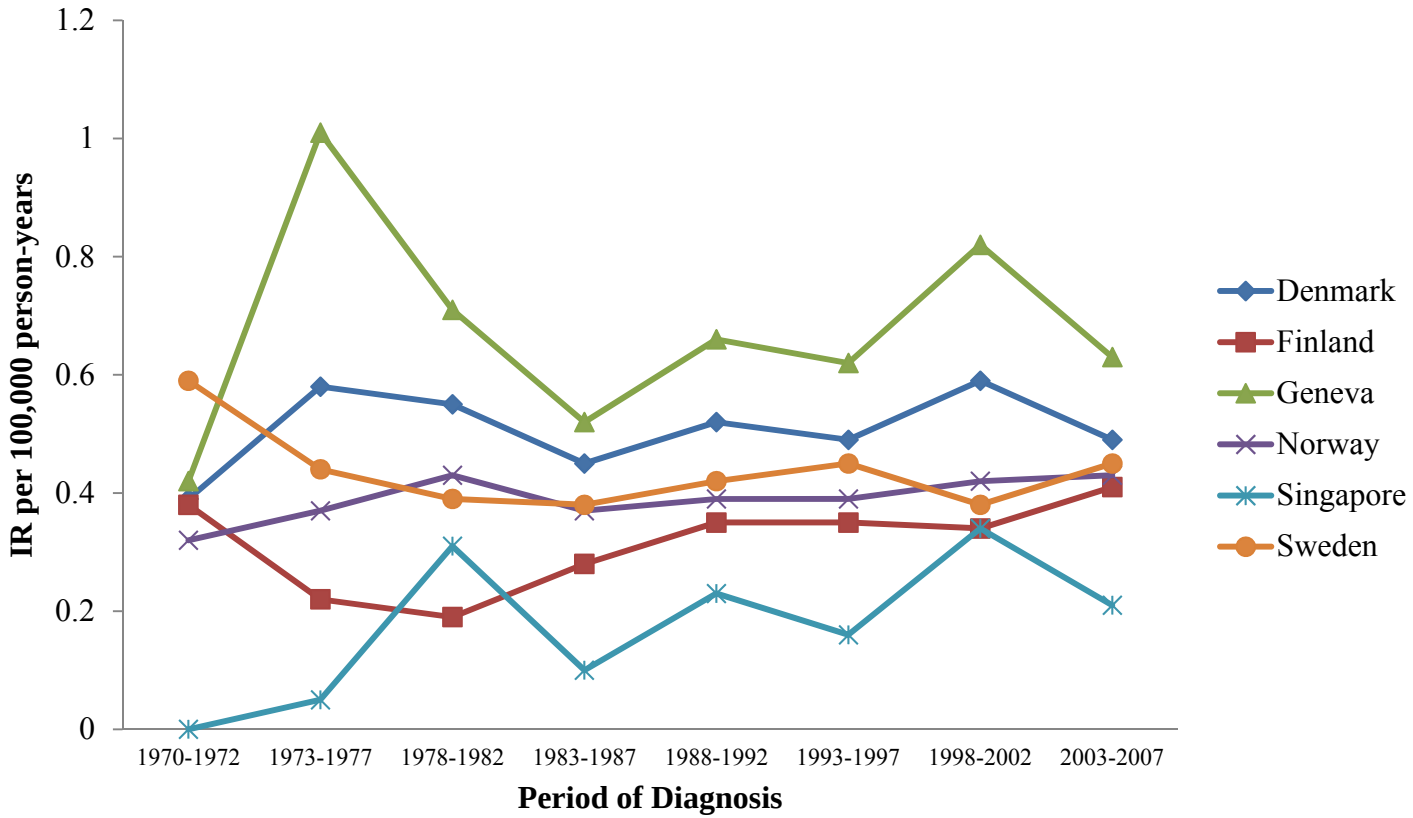


Figure 5.3 Incidence rate (IR) of female invasive breast cancer (standardized to world population), by period of diagnosis and region

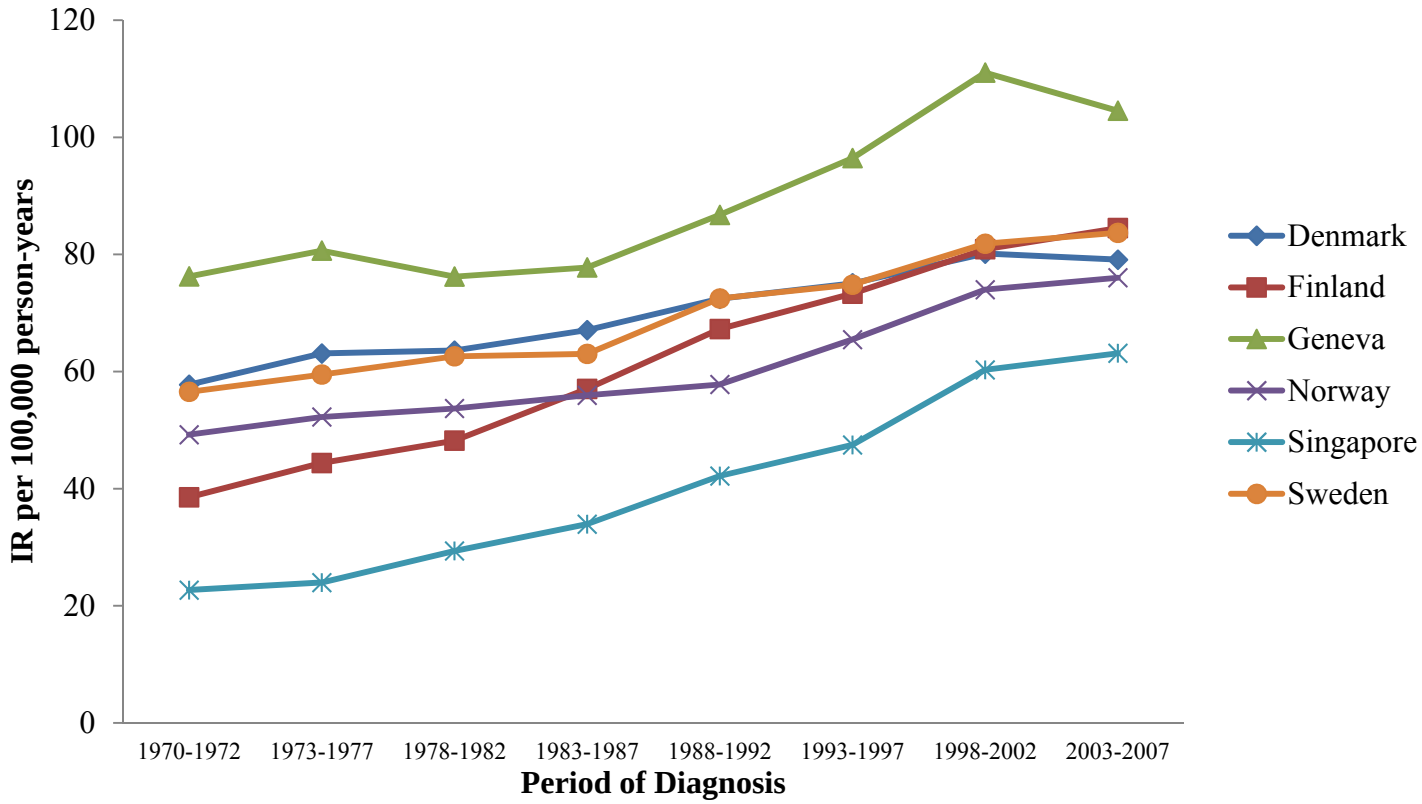
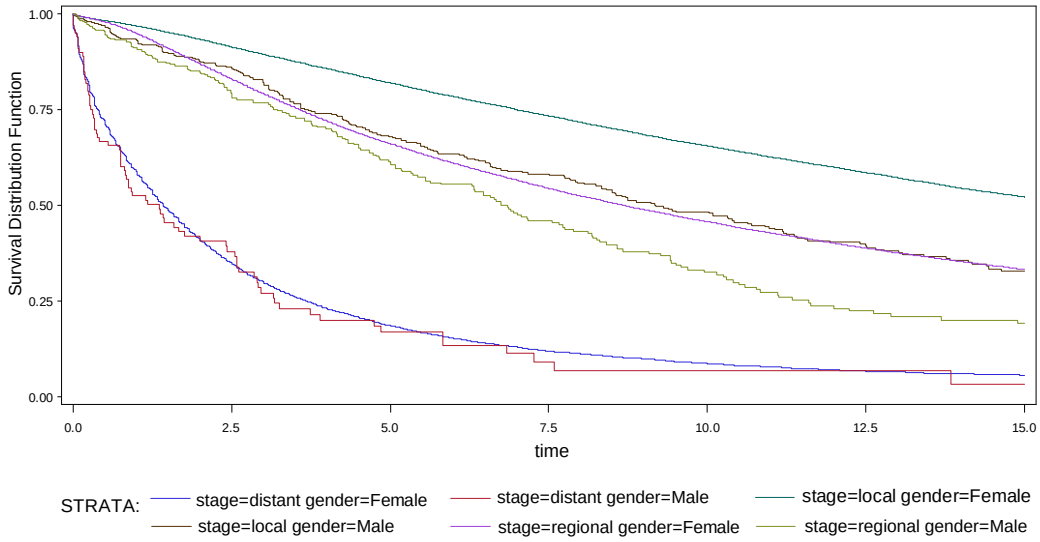


Figure 5.4 Kaplan-Meier curve for overall survival of breast cancer patients, by gender and stage



Male

TIME (years)	0	5	10	15
Localized	374	218	127	57
Regional	311	136	50	23
Distant	100	12	4	1

Female

TIME (years)	0	5	10	15
Localized	83457	54831	32616	18394
Regional	64522	31700	15572	7770
Distant	10526	1415	498	238

Table 5.2 The 5-year and 15-year relative survival ratios (RSRs) and relative excess risk differences between male and female breast cancer patients

Sex	RSR (95% CI)	Relative excess risk (95% CI)						
		Model 1a	**Model 1b**	**Model 2a**	**Model 2b**	**Model 3a**	**Model 3b**	**Model 4**
5-year follow-up								
Male	0.72 (0.70,0.75)	1.27 (1.13,1.42)	1.20 (0.97,1.50)	1.13 (1.01,1.26)	1.08 (0.88,1.33)	0.96 (0.80,1.15)	0.91 (0.75,1.12)	0.78 (0.62,0.97)
Female	0.78 (0.78,0.78)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)
15-year follow-up								
Male	0.50 (0.46,0.54)	1.36 (1.24,1.50)	1.24 (1.03,1.49)	1.16 (1.06,1.28)	1.09 (0.91,1.30)	1.00 (0.85,1.18)	0.96 (0.81,1.15)	0.80 (0.65,0.98)
Female	0.61 (0.60,0.61)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)

RSR, Model 1a and 1b:	Crude
Model 2a and 2b:	Adjust for region, time since diagnosis, age and year of diagnosis
Model 3a and 3b:	Adjust for region, time since diagnosis, age and year of diagnosis and stage
Model 4:	Adjust for region, time since diagnosis, age and year of diagnosis, stage and treatment (surgery, chemotherapy, radiotherapy and hormonal therapy)
RSR, Model 1a and 2a:	Entire dataset, N=462,511
Model 3a:	Subset with information on stage (Finland, Geneva, Norway and Singapore), N=190,030
Model 1b, 2b,3b and 4	Subset with information on treatment (Finland, Geneva and Norway), N=167,169

Table 5.3 Breast cancer 5-year relative survival ratio, by gender, calendar period and stage

Calendar period	Relative survival ratio (95% CI)			
	Overall*	**Localized**	**Regional**	**Distant**
		Male (N=917)		
1970-1977	0.67 (0.60,0.72)	0.81 (0.65,0.95)	0.70 (0.47,0.90)	0.16 (0.03,0.43)
1978-1987	0.68 (0.63,0.73)	0.85 (0.71,0.97)	0.68 (0.51,0.82)	0.26 (0.06,0.57)
1988-1997	0.73 (0.68,0.78)	0.97 (0.86,1.05)	0.72 (0.56,0.85)	0.08 (0.01,0.22)
1998-2007	0.78 (0.73,0.83)	0.81 (0.65,0.93)	0.92 (0.80,1.01)	0.39 (0.17,0.63)
Overall	0.72 (0.70,0.75)	0.87 (0.81,0.93)	0.78 (0.71,0.85)	0.23 (0.13,0.34)
		Female (N=189,113)		
1970-1977	0.66 (0.66,0.67)	0.84 (0.83,0.85)	0.54 (0.53,0.56)	0.13 (0.12,0.15)
1978-1987	0.73 (0.73,0.74)	0.89 (0.88,0.90)	0.65 (0.64,0.66)	0.17 (0.15,0.19)
1988-1997	0.80 (0.80,0.80)	0.93 (0.92,0.93)	0.74 (0.74,0.75)	0.22 (0.20,0.24)
1998-2007	0.87 (0.86,0.87)	0.97 (0.96,0.97)	0.86 (0.85,0.86)	0.31 (0.29,0.33)
Overall	0.78 (0.78,0.78)	0.92 (0.92,0.92)	0.73 (0.73,0.74)	0.22 (0.21,0.23)

* Estimated using entire dataset, N=462,511

5.4 Discussion

With this international population-based study, we show that over the last 38 years, male breast cancer incidence has remained at a stable low rate, whereas female breast cancer has become increasingly common. In a crude comparison survival is worse among men than among women. The poorer observed survival of male patients is largely explained by their more advanced stage at diagnosis, their higher age at diagnosis, and lower proportion being treated with locoregional treatment. After adjusting for these factors, men actually had better relative survival than women.

Over the last 40 years, the risk of breast cancer in women has continued to increase at a steady pace, largely explained by the introduction of mammography screening and hormone replacement therapy in the 1980s [43]. Additionally, changes in lifestyle and reproductive patterns (i.e. age at menarche, age at first birth, parity, and frequency and duration of breast feeding) have influenced female breast cancer risks. Virtually all of these factors, except changes in lifestyle, have not affected men over time.

Breast cancer is diagnosed on average 5 to 10 years later in men than in women [11, 274, 278, 279]. Because of the lack of early detection by mammography and awareness of early signs of breast cancer, the duration of symptoms before diagnosis has been reported to be longer in men, with a median of 4 to 6 months [274, 282]. This may contribute to the differences in stage distribution between men and women. In our study, the proportion of distant spread of disease stage was two-fold in men compared with women.

Female breast cancer patients have experienced substantial improvements in survival over the last 30 years. The improvement in male breast cancer survival is not as pronounced. The survival improvement in women is partly explained by the introduction of screening (both opportunistic and within national programmes targeted to women only), leading to earlier detection and detection of indolent tumors and over-diagnosis. Advances in treatment (in particular, the introduction of tamoxifen in the 1980s) and standardization of treatment regimens in international guidelines have improved breast cancer survival probabilities [7, 147].

Lack of evidence-based treatment guidelines and differences in compliance with treatment may explain why men experience less survival benefit than women. Locoregional and adjuvant treatment of male breast cancer has not yet been evaluated in randomized trials, and evidence-based treatment guidelines are lacking. As a result, most clinicians base their treatment strategy on guidelines for female breast cancer. However, systemic treatment, especially anti-hormonal treatment is not as straightforward in men. Anti-estrogen treatments like tamoxifen are not well tolerated by men, resulting in lower treamtent compliance [283-285]. In addition, National Comprehensive Cancer Network guidelines suggest that administration of aromatase inhibitors to men is not effective without simultaneous suppression of testicular steroidogenesis [286].

Although the observed survival of male breast cancer is worse than that of female disease, male gender is not an independent risk factor of poor outcome after breast cancer. Actually, our results suggest the opposite, that male gender is a favourable prognostic factor, as shown by the reduced relative excess risk of death after breast cancer, after adjustment for age at diagnosis, stage and treatment. This is in line

with other studies [126, 280] that found a similar relative survival for women and men with early-stage breast cancer, whereas men with late-stage breast cancer had better survival than women. Our stepwise adjustment shows that stage and treatment differences between female and male patients explain most of the poorer (unadjusted) relative survival of men. These results suggest that much improvement in outcome of male breast cancer can be achieved by improving earlier detection (through awareness and promotion of breast self-examination) and development of treatment guidelines.

We acknowledge that our study suffers from limitations. An unavoidable limitation of a study with a time frame of almost 40 years involves the improvements in diagnostic performance and the increased diagnostic intensity in women, which has led to an increased uptake of small, often indolent cancers. Other limitations are discrepancies in staging system among registries and lack of information on tumor characteristics such as grade, hormone receptor status and details on systemic treatment.

Strengths of our study include the large number of male patients, the long observation time and the high-quality (population-based) data and the completeness of follow-up, which allow for unbiased ascertainment of cancers and deaths. Unlike previous studies that compared outcome in male and female breast cancer [276, 287], we accounted for gender differences in life expectancy by looking at relative survival and relative excess risk.

In conclusion, male breast cancer risk has remained constant over the last 40 years. Male patients have later onset and more advanced disease than female patients. Overall survival of male breast cancer is worse, however after adjustment of life

expectancy, age and year of diagnosis, stage and treatment, male breast cancer patients actually emerged as having a survival benefit compared with women.

Chapter 6 The impact of *in situ* breast cancer and family history on risk of subsequent breast cancer events and mortality: a population-based study from Sweden

6.1 Motivation

Women with *in situ* breast cancer have an increased risk of developing *in situ* or invasive breast cancer in the ipsilateral or contralateral breast [169, 288-297]. Moreover women with *in situ* breast cancer, even after treatment, are at increased risk of subsequent invasive breast cancer compared to women in the general population [169, 249, 288, 290-295, 298, 299]. The clinical behavior of *in situ* breast cancer is incompletely understood but it is likely that it represents a mixed population of indolent and more aggressive tumors. Several factors have been associated with invasive recurrences, including patient characteristics [169, 291, 294], tumor characteristics [169, 291, 300][4, 5, 15] and treatment [168, 169, 301]. The influence of a positive family history on subsequent breast cancer is less well studied [302-304].

The risk of death from breast cancer in women diagnosed with *in situ* breast cancer is considered to be at most only marginally increased, but remains less well characterized and, with few exceptions, studies are often limited by short follow-up and non-population-based designs [249, 305].

In this study we evaluated the long-term risk of second breast cancer and death among women diagnosed with *in situ* breast cancer, in relation to family history.

6.2 Methods

We selected 8111 women from Swedish Multi-Generation register (MGR) who were diagnosed with breast carcinoma *in situ* between 1st January 1980 and 1st January 2005. Information on family history, any subsequent invasive breast cancer or *in situ* disease in the contralateral breast after the first diagnosis was retrieved from the cancer registry. Family history was defined as having at least one first-degree relative diagnosed with invasive breast cancer at any point in time.

Any invasive cancer following *in situ* breast cancer was reported as a new event, as were new *in situ* breast cancers in the contralateral breast. Local relapses were not recorded. Ipsilateral *in situ* breast cancer was excluded due to the increased probability of being underreported in women with previous *in situ* breast cancer. Thus, we defined subsequent breast events as ipsilateral or contralateral invasive or a contralateral *in situ* breast cancer. Women with any previous invasive or *in situ* breast cancer were excluded, as were women with invasive breast cancer diagnosed concurrently with the first *in situ* breast cancer. Because incomplete information on laterality and *in situ* breast cancer registration prior to 1980, we restricted our cohort to women with a first *in situ* breast cancer diagnosed after 1980. Linkage with the Cause of Death Register and the Total Population Register provided us follow-up information regards to death, immigration and emigration.

Statistical analysis

To estimate the risk of a subsequent breast event (ipsilateral or contralateral invasive or a contralateral *in situ* breast cancer), all women were followed from the date of their first *in situ* breast cancer diagnosis and continued until a subsequent breast cancer, emigration, death, or end of follow-up, whichever came first. We

estimated standardized incidence ratios (SIRs) as a measure of relative risk. The expected number of subsequent breast cancer events was calculated as the product of the person-years accumulated by women with *in situ* breast cancer by the age- and calendar period-specific incidence of unilateral *in situ* /invasive breast cancer of the female population in the MGR. Thus SIRs compare gender, age- and calendar-adjusted risk of subsequent breast events of *in situ* breast cancer patients to that of the general population. For all estimates of the contralateral breast, the background rate of *in situ* and invasive breast cancer was divided by two, as only one breast was "at risk". SIRs of subsequent invasive breast cancer was calculated for calendar period of first diagnosis, age and time since first diagnosis and stratified by family history of breast cancer. Poisson trend tests for monotonic trend of SIR across calendar period, age and time since first diagnosis was performed [261]. We used Poisson regression modeling among women with a first *in situ* breast cancer to estimate the independent effects of age, year of diagnosis and time since diagnosis as well as effect of family history on the risk of ipsilateral or contralateral invasive or contralateral *in situ* breast cancer. Since background rates of breast cancer vary considerably by age we also estimated excess additive risks (EARs), as the difference of observed numbers of subsequent invasive breast cancer and the expected number in the general population in the Swedish MGR, as a measure of absolute risk for subsequent invasive cancer. EARs were estimated using a univariate Poisson model with an identity link function and the expected number of cases as the offset. A likelihood ratio test was used to calculate 95% CIs. The cumulative incidence was estimated using life table (actuarial) method.

The standardized mortality ratio (SMR) was used as a measure of relative mortality. The expected number of deaths was calculated from the general population in the

MGR. SMRs were also stratified by family history, age at first *in situ* breast cancer diagnosis and type of subsequent breast event. For overall SMRs, subjects were followed from the date of first *in situ* breast cancer diagnosis until date of emigration, death, or end of follow-up, whichever came first. In contrast, in the estimates of death by type of subsequent breast event, follow up was started at the diagnosis of that particular event. We calculated 95% CIs assuming a Poisson distribution for the observed number of cases. All data preparation and analysis was done using the SAS statistical package, version 8.2 or higher (SAS Institute Inc., Cary, NC, USA).

6.3 Results

Patient characteristics are listed in Table 6.1. Over a follow-up period of 71,458 person-years, 825 (10.2%) women developed 886 subsequent breast events (118 contralateral *in situ* and 768 ipsilateral or contralateral invasive breast cancers). The proportion of subsequent breast events was similar in women with as well as without a family history (11.3%, n=97 versus 10.0%, n=728). The average time from first *in situ* breast cancer diagnosis to a second breast event was overall 5.6 years.

Table 6.1 Summary of all women diagnosed with *in situ* breast cancer from 1980 to 2004.

		All	No family history	Family history
Total		8111	7252	859
Mean age at first *in situ* breast cancer (SD)		59.09 (12.1)	59.7 (12.1)	53.9 (10.8)
Mean follow-up time, years (SD)		8.8 (5.9)	8.3 (5.9)	7.7 (5.4)
Year at diagnosis of first *in situ*	1980-1984	665	624	41
	1985-1989	1211	1108	103
	1990-1994	2046	1835	211
	1995-1999	1963	1727	236
	2000-2004	2226	1958	268
Age at diagnosis of first *in situ*	Less than 40	335	269	66
	40-44	594	507	87
	45-49	1078	903	175
	50-54	1313	1133	180
	55-59	1133	993	140
	60-64	1021	943	78
	65-69	1058	995	63
	70-74	778	748	30
	75+	801	761	40

Type of second events	Contralat *in situ*[1]	118	104	14
	Ipsilat invasive	376	334	42
	Contralat invasive	303	262	41
	Total invasive[1]	768	677	91
	Second breast event total[1,2]	886	781	105
Type of second events (# of women)	Contralat *in situ* [1]	117	103	14
	Ipsilat invasive	370	328	42
	Contralat invasive	299	258	41
	Total invasive[1]	725	637	88
	Second breast event total[1,2]	825	728	97

[1]includes the events where laterality is missing.

[2]ipsilateral *in situ* events is not included in the study

Risk of subsequent breast cancer/in situ

Table 6.2 presents the risk of second invasive or *in situ* breast cancer. The risk of a subsequent ipsilateral or contralateral invasive breast cancer was increased more than fourfold (SIR=4.55, 95% CI, 4.23- 4.88) among women with *in situ* breast cancer as compared to women in the general population and the risk for a contralateral *in situ* breast cancer was almost sixteenfold increased (SIR=15.98, 95% CI, 13.23-19.14). Poisson regression analyses showed that women with a family history of breast cancer had almost 50 percent increased risk of contralateral invasive breast cancer, compared to women without a family history of breast cancer (adjusted incidence rate ratio =1.47, 95% CI, 1.05-2.05).

Among women diagnosed with *in situ* breast cancer, the cumulative 10- and 20-year risk for a subsequent ipsilateral or contralateral invasive cancer was approximately 10 and 18 percent respectively, while the cumulative 10- and 20-year risk for a subsequent contralateral *in situ* breast cancer was 1 and 2 percent respectively (Figure 6.1).

Table 6.2 Standardized incidence ratio (SIR) of second breast event (contralateral *in situ* or ipsilateral or contralateral invasive breast cancers) after diagnosis of first *in situ* breast cancer and its 95% CI, by type of second breast event and family history.

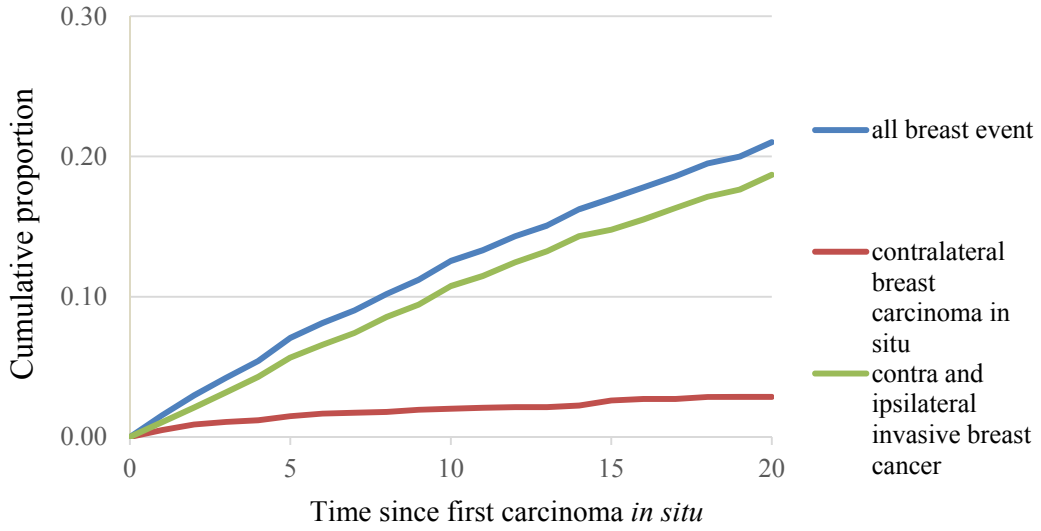
	All		**No family history**		**Family history**		**Incidence Rate Ratio* (95% CI)**
	No. of cases	SIR (95 % CI)	No. of cases	SIR (95 % CI)	No. of cases	SIR (95 % CI)	
2nd breast cancer†	886	5.08 (4.75,5.43)	781	4.95 (4.61,5.31)	105	6.27 (5.13,7.60)	1.17 (0.95,1.44)
2nd *in situ* Contralateral†	118	15.98 (13.23,19.14)	104	15.80 (12.91,19.15)	14	17.44 (9.54,29.26)	1.09 (0.62,1.92)
2nd invasive Ipsilateral +contralateral +missing side	768	4.55 (4.23,4.88)	677	4.40 (4.07,4.74)	91	5.62 (4.53,6.91)	1.19 (0.95,1.49)
2nd ipsilateral invasive‡	376	4.26 (3.84,4.72)	334	4.19 (3.75,4.66)	42	4.97 (3.58,6.72)	1.00 (0.72,1.38)
2nd contralateral invasive‡	303	3.42 (3.05, 3.83)	262	3.28 (2.89,3.70)	41	4.82 (3.46,6.54)	1.47 (1.05, 2.05)

* Reference group is No family History. Incidence rate ratio has been adjusted for attend age, calendar period, age and year of first diagnosis of carcinoma *in situ* and time since first diagnosis

† background rate of *in situ* breast cancer was divided by 2

‡ background rate of invasive breast cancer was divided by 2

Figure 6.1 Cumulative incidence of second breast event among women diagnosed with *in situ* breast cancer, stratified by types of subsequent breast events.



Women with *in situ* breast cancer with no family history experienced an increasing SIR of a subsequent invasive cancer during the study period, SIR 3.09 (95% CI, 2.42-3.89) in 1980-1984, versus SIR 5.05 (95% CI, 3.88-6.46) in 2000-2004 (P-trend <0.001). In contrast, for women with a family history, SIR of a subsequent invasive breast cancer remained relatively high over the study period (Table 6.3). The EAR also increased over the study period for women with no family history but not for women with a family history (Table 6.4).

Overall, the relative risk for a subsequent invasive breast cancer was almost twice as high for women under forty at first *in situ* breast cancer diagnosis compared with women over forty, SIR 8.54 (95% CI, 6.07-11.67) and 4.44 (95% CI, 4.12-4.77) respectively (P-value <0.001). Among women below forty with a positive family history, the risk for a subsequent invasive cancer was more than fourteen times higher than in the general population, SIR 14.3 (95% CI, 7.39-24.99). Given that the background rates of breast cancer are highly age-dependent, we estimated the

EAR in relation to age at diagnosis. While the relative risk of a subsequent invasive breast event decreased with increasing age for both women with and without a family history for breast cancer, the overall EAR was similar for women below forty years at diagnosis (93.17 per 10,000 person-years, 95% CI, 63.42-129.84) as compared to women over forty (88.50 per 10,000 person-years, 95% CI, 80.41-96.99) (Table 6.4). In contrast, women with a family history of breast cancer had higher EAR, with women under 40 years of age carrying the greatest EAR (154.10 per 10,000 person-years, 95% CI, 77.14-266.30), compared to women older than 40 years at diagnosis (105.72 per 10,000 person-years, 95% CI, 78.88-136.82). This suggests that both relative and absolute risks are higher with younger age of onset of *in situ* disease in women with a positive family history.

Finally, regardless of family history, the risk for subsequent invasive cancer in the first five years after first *in situ* breast cancer was increased more than fivefold compared to the general population (SIR=5.20, 95% CI, 4.71-5.74). In women with no family history there was a significant decline in both the relative and absolute risk over time, but this was not observed in women with a family history (Table 6.3 and 6.4).

Table 6.3 Standardized incidence ratio (SIR) of second invasive breast cancer (ipsilateral and contralateral) after diagnosis of first *in situ* breast cancer, by year at first diagnosis, age at first diagnosis, time since first diagnosis and family history

		Overall		**No Family History**		**Family History**	
		No. of cases	SIR (95% CI)	No. of cases	SIR (95% CI)	No. of cases	SIR (95% CI)
Calendar year*	1980-1984	81	3.31 (2.63,4.11)	72	3.09 (2.42,3.89)	9	6.58 (3.01,12.49)
	1985-1989	141	3.54 (2.98,4.17)	123	3.36 (2.79 ,4.01)	18	5.34 (3.16 ,8.44)
	1990-1994	292	5.23 (4.65,5.86)	259	5.12 (4.51,5.78)	33	5.90 (4.06,8.29)
	1995-1999	182	5.24 (4.51,6.06)	160	5.17 (4.40,6.04)	22	5.23 (3.28,7.92)
	2000-2004	72	5.11 (4.00,6.44)	63	5.05 (3.88 ,6.46)	9	5.48 (2.51,10.40)
	P-trend		<0.001		<0.001		1
Age at diagnosis	< 40	39	8.54(6.07,11.67)	27	7.20(4.75,10.48)	12	14.30 (7.39,24.99)
	40-49	173	4.88 (4.18,5.66)	147	4.85 (4.10,5.71)	26	4.70 (3.07,6.89)
	50-59	221	4.07 (3.55,4.65)	189	3.88 (3.35,4.48)	32	5.22 (3.57,7.37)
	60-69	220	4.57 (3.99,5.22)	207	4.53 (3.93,5.19)	13	5.16 (2.75,8.82)
	≥70	115	4.33 (3.58,5.20)	107	4.20 (3.45,5.08)	8	6.90 (2.98,13.60)
	P-trend		0.008		0.069		0.096
	<40	39	8.54(6.07,11.67)	27	7.20(4.75,10.48)	12	14.30 (7.39,24.99)
	>40	729	4.44 (4.12,4.77)	650	4.33 (4.00 ,4.67)	79	5.15 (4.08,6.42)
	P-value		<0.001		0.012		0.001
Time since diagnosis	0-4	401	5.20 (4.71,5.74)	359	5.13 (4.62,5.69)	42	5.45 (3.93,7.37)
	5-9	230	4.44 (3.89,5.06)	197	4.19 (3.62,4.82)	33	6.51 (4.48,9.15)
	10-14	96	3.42 (2.77,4.17)	85	3.31 (2.65,4.10)	11	4.28 (2.14,7.66)
	15+	41	3.41 (2.44,4.62)	36	3.19 (2.24,4.42)	5	5.92 (1.92,13.82)
	P-trend		<0.001		<0.001		0.848

* When the follow-up time was restricted to 5 years the estimates were similar but the trend tests not significant

Table 6.4 Excess additive risk (EAR) of second invasive breast cancer (ipsilateral and contralateral) per 10,000 person-years after diagnosis of first *in situ* breast cancer, by year at first diagnosis, age at first diagnosis, time since first diagnosis and family history

		Overall		**No Family History**		**Family History**	
		No. of cases	EAR (95% CI)	No. of cases	EAR (95% CI)	No. of cases	EAR (95% CI)
Calendar year	1980-1984	81	52.75 (37.45,70.43)	72	48.28 (33.04,66.05)	9	110.89 (42.95,215.85)
	1985-1989	141	61.81 (48.36,76.82)	123	57.73 (44.05,73.13)	18	100.78 (51.95,167.20)
	1990-1994	292	107.88 (93.16,123.77)	259	105.83 (90.45,122.51)	33	116.53 (73.94,169.99)
	1995-1999	182	111.20 (92.24,132.08)	160	109.53 (89.57,131.67)	22	110.38 (61.01,175.62)
	2000-2004	72	109.35 (80.32,143.22)	63	108.33 (77.66,144.49)	9	114.67 (41.80,227.26)
Age at diagnosis	< 40	39	93.17 (63.42,129.84)	27	77.90 (47.93,116.43)	12	154.10 (77.14,266.30)
	40-49	173	85.36 (70.19,102.11)	147	84.82 (68.47,103.03)	26	81.46 (46.61,126.48)
	50-59	221	83.61 (69.64,98.88)	189	78.20 (63.88, 93.94)	32	117.94 (73.05,174.48)
	60-69	220	98.46 (82.54,115.85)	207	96.87 (80.69,114.59)	13	121.1 (53.53,218.20)
	≥70	115	85.99 (66.78,107.69)	107	82.72 (63.43,104.60)	8	152.17 (55.54,305.46)
Time since diagnosis	<40	39	93.17 (63.42,129.84)	27	77.90 (47.93,116.43)	12	154.10 (77.14,266.30)
	>40	729	88.50 (80.41,96.99)	650	85.87 (77.50,94.67)	79	105.72 (78.88,136.82)
	0-4	401	98.39 (86.87,110.69)	359	97.42 (85.34,110.38)	42	98.33 (65.49,138.51)
	5-9	230	88.49 (74.36,103.89)	197	82.22 (67.83, 98.01)	33	137.92 (88.46,200.00)
	10-14	96	66.28 (48.77,86.29)	85	63.44 (45.47,84.15)	11	89.41 (33.33,172.54)
	15+	41	69.06 (42.11,102.12)	36	62.90 (36.15,96.16)	5	143.72 (32.81,342.47)

Mortality risk

The overall all risk of death in women with *in situ* was significantly increased, by 30 percent compared to the general population but highly dependent on the occurrence of second invasive cancer event (Table 6.5). Women, who did not develop a second invasive event following *in situ* breast cancer, had a similar risk of death as the background population (SMR=1.01, 95% CI, 0.95-1.08). In contrast, women who were diagnosed with a second invasive event were twice as likely to die as compared to women in the general population (SMR=2.06, 95% CI, 1.72-2.44) with no significant differences between women with and without a family history for breast cancer.

The overall risk of death following an *in situ* breast cancer was increased for women with a family history (SMR=1.44, 95% CI, 1.15-1.78) as well as for women without (SMR=1.28, 95% CI, 1.21-1.35). Given that deaths were rare at younger ages we compared mortality among women above and below age 50 years. Women below age 50 years at first *in situ* breast cancer diagnosis and who were diagnosed with a second invasive cancer, had significantly higher mortality as compared to women over 50 years at diagnosis, (SMR=8.03, 95% CI 5.38-11.54 versus SMR=1.70, 95% CI 1.39-2.06). The laterality of the second invasive event did not influence the risk of death.

Table 6.5 Standardized mortality ratio (SMR) of second breast event (contralateral *in situ* or ipsilateral or contralateral invasive breast cancers) after diagnosis of first *in situ* breast cancer and its 95% CI, by type of second breast event and family history.

	All		**No family history**		**Family history**		**<50**		**>50**	
	No. of deaths	SMR (95 % CI)	No. of deaths	SMR (95 % CI)	No. of deaths	SMR (95 % CI)	No. of deaths	SMR (95 % CI)	No. of deaths	SMR (95 % CI)
Overall	1343	1.28 (1.22,1.36)	1258	1.28 (1.21,1.35)	85	1.44 (1.15,1.78)	122	2.19 (1.82,2.61)	1221	1.24 (1.17,1.31)
No event + 2nd contralateral *in situ*	927	1.01 (0.95,1.08)	875	1.01 (0.94,1.08)	52	1.02 (0.76,1.34)	58	1.17 (0.89,1.52)	869	1.00 (0.93,1.07)
2nd invasive ipsilateral +contralateral +missing side	132	2.06 (1.72,2.44)	122	2.03 (1.68,2.42)	10	2.54 (1.22,4.67)	29	8.03 (5.38,11.54)	103	1.70 (1.39,2.06)
2nd ipsilateral invasive	63	2.16 (1.66,2.77)	58	2.12 (1.61,2.74)	5*	2.75 (0.89,6.43)	17	12.89 (7.51,20.64)	46	1.65 (1.21,2.21)
2ndcontralateral invasive	55	1.99 (1.50,2.59)	49	1.92 (1.42,2.54)	6*	2.82 (1.04,6.15)	10	7.85 (3.77,14.44)	45	1.71 (1.24,2.28)

*one subject had both ipsilateral and contralateral invasive breast cancer

6.4 Discussion

In this large population-based cohort, with data from nationwide, high quality registers, we demonstrated that women diagnosed with *in situ* breast cancer had a considerably increased risk for an invasive and contralateral *in situ* breast cancer, compared to women in the general population, with young women facing the highest risks. Having a positive family history increased the risk for a contralateral invasive breast cancer by 50 percent compared to not having a family history for breast cancer. The increased risk for an invasive cancer persisted over time and still fifteen years after diagnosis, the risk was three times higher than in women in the general population. Meanwhile, the mortality for women with *in situ* breast cancer was the same as the general population, as long as an invasive cancer did not occur.

In women diagnosed with breast carcinoma *in situ* and with a positive family history, the risk of a contralateral invasive breast cancer was four times higher than women in the general population. It is 1.5 times higher compared to women with *in situ* disease but without family history. There are methodological issues that may account for these differences, since our estimates assume only one breast is at risk, with a corresponding lower expected rate. Two meta-analyses of familiar risks for breast cancer presented the relative risk associated with having a first degree relative of breast cancer to be 2.1 and 1.8, respectively [28, 306]. The observed diluted additional risk in women with a family history, i.e. only 50 percent increased risk for a contralateral invasive cancer, and no increased risk for ipsilateral invasive cancer or contralateral *in situ* cancer, as compared to non-family history women, is intriguing. We speculate that women with a positive family history were likely more prone to choose mastectomy than those without family history, which would

reduce the risk for an ipsilateral cancer in these women. The reduced risk may also be a reflection of heterogeneity of the *in situ* breast cancer phenotype. Additional stratification into one, two or even three affected first-degree members to better quantify the hereditary component may have allowed a deeper understanding of these results.

Regardless of family history, women under forty years of age at diagnosis had a significantly higher risk for subsequent invasive breast cancer compared to women above forty years. These young women would experience an absolute excess risk ranging from about 8 events per 1,000 person-years to as high as 15 events per 1,000 person-years depending on family history, this absolute excess risk decreases with increasing age only for women with a positive family history. Given that young women with family history have much higher risk for a subsequent event, studies on best treatment options such as mastectomy versus breast conserving surgery may be needed for this group of patients.

The increased relative risk for subsequent invasive breast cancer by almost 60 percent from 1980-84 to 2000-04, exclusively in women with no family history, may be related to a combination of screening and treatment patterns. During the study period, nationwide mammography screening was introduced, which had a complete national coverage by 1997 [307]. With increasing mammography screening and subsequently a larger number of detected smaller lesions, the majority of whom are non-palpable, the use of breast-conserving surgery has become the norm from 1990 onwards [308]. In comparison to mastectomy, breast-conserving surgery poses an increased risk for both local recurrence and new ipsilateral primary cancers. In contrast, women with a positive family history had

no increased risk during the study period and we speculate that these women, who had relatives with breast cancer, were more prone to choose mastectomy.

During follow-up, women with no family history of breast cancer had a gradually decreasing risk for subsequent invasive breast cancer with time since diagnosis. However, still 15 years after first *in situ* breast cancer, the risk for an invasive breast cancer was almost three times higher than for women in the general population. This indicates that women diagnosed with *in situ* breast cancer have a lifelong increased risk, which needs to be taken into account for when planning their follow-up.

Overall, there was no increased risk of death for women with *in situ* breast cancer as long as a second invasive event did not occur, but in women with a second invasive breast cancer the risk of death was doubled. There were no significant differences in mortality between women with and without family history. Young age of onset was an important predictor of death for women with *in situ* disease due to an increased risk for second invasive cancers and thus a substantially higher mortality, which should be taken into account when planning their treatment and follow-up.

Strengths of the current study include the population-based design, its large sample size, complete follow-up and unbiased ascertainment of family history, cancers and death. To the best of our knowledge, this is the largest study to assess the impact of a positive family history for breast cancer on risk and mortality after *in situ* breast cancer.

This study has a number of limitations. We have not distinguished between mastectomies and breast-conserving surgery, and ductal carcinoma *in situ* breast

cancer and lobular carcinoma *in situ* breast cancer. With this stated, a previous Swedish case-control study has shown that the risk for a subsequent invasive breast cancer was equal after lobular and ductal carcinoma *in situ* breast cancer [300]. Due to regional differences in how to report second ipsilateral *in situ* breast cancer, such events were not included in the study.

In conclusion, a positive family history increases the risk only for a contralateral invasive breast cancer among women with *in situ* breast cancer. The risk for a subsequent invasive breast cancer, as well as mortality is substantially higher in younger women, which should be taken into account when planning their treatment and follow-up.

Chapter 7 Validation of the CancerMath prognostic tool for breast cancer in Southeast Asia

7.1 Motivation

Adjuvant chemotherapy and hormone therapy improve long-term survival and reduce the risk of recurrence in early breast cancer patients [147, 309, 310]. However, the benefit varies greatly from patient to patient due to biologic heterogeneity of the disease and differences in response to treatment [164, 311]. Risk of adverse effects and high cost of adjuvant therapy also make it challenging for oncologists to choose the most appropriate treatment. Therefore, several clinical tools have been developed to predict prognosis and survival benefit from treatment, using clinical and histological features, genetic profiles, and novel biomarkers [185].

CancerMath (http://www.lifemath.net/cancer/) is the latest web-based calculator to estimate overall survival for each of the first 15 years after diagnosis [189]. It also provides information on conditional survival (the likelihood of surviving given being alive after a certain number of years), probability of positive lymph nodes, and nipple involvement as well as benefit of systemic treatment. However this new tool has not been validated outside the United States. The aim of the study is to validate this model in the Singapore Malaysia Hospital Based Breast Cancer Registry, demonstrating its predictive performance for different subgroups and determining its calibration and discrimination.

7.2 Methods

Women diagnosed with pathological stage I to III breast cancer according to the sixth edition AJCC staging system, who underwent surgery, were identified from the Singapore Malaysia Hospital Based Breast Cancer Registry. Patients diagnosed until 31st December 2011 were followed up from date of diagnosis until date of death or date of last fellow-up, whichever came first. Date of last follow-up was 1st March 2013 for UMMC, 31st July 2013 for NUH, and 1st October 2012 for TTSH. Male patients, patients with unknown age at diagnosis and tumor size were excluded from this analysis as these two were essential predictors for all four CancerMath calculators.

Javascript code of all four CancerMath calculators which contained predetermined parameters for the predictors and mathematical equations was exported on 9th Nov 2013 from its website by selecting "view-> source" in the browser menu. The script was then transcribed into R script to allow calculation for a group of patients. For nodal status calculator, patient's age, tumor size, ER and PR status, histological type, and grade were used by the programme to calculate probability of positive nodes for each patient. Overall mortality risk at each year up to 15 year after diagnoses was predicted by outcome calculator, based on age, tumor size, number of positive nodes, grade, histological type, ER, PR, and HER2 status. Effect of hormone and chemotherapeutic regimen on overall mortality was further adjusted by the therapy calculator and number of years since diagnosis was taken into account in conditional survival calculator. Results from R script and website were crosschecked with a random subset of 20 patients to verify the accuracy of the R script. Histological type recorded as others was re-categorized as unknown

histological type for calculation. If HER2 status was equivocal based on IHC and FISH was not performed, HER2 status was treated as unknown. Evidence of recurrence was set as unknown for conditional survival calculation.

Only cases with known nodal status (N=6807) were included for validation of nodal status calculator and their individual probability of positive lymph nodes was calculated. For outcome calculator, two separate subsets of patients with minimum 5-year follow up (UMMC and NUH patients diagnosed in 2007 and earlier and TTSH patient diagnosed in 2006 and earlier, N=4517) and patients with 10-year follow-up (UMMC and NUH cases diagnosed in 2002 and earlier, N=1649) were selected for comparison of observed and predicted survival. As NUH and TTSH did not collect details of hormone therapy and chemotherapy regimen data, therapy calculator was only validated for UMMC patients with minimum 5-year follow up (N=1538).

Statistical analysis

Nodal status calculator

Observed and predicted probabilities of positive lymph nodes were compared. Calibration was assessed by plotting the observed probability of positive nodes against median of predicted probability for each decile of the predicted probability. Discrimination of predicted probability of nodal involvement was evaluated by AUC.

Outcome and therapy calculator

Ratio of observed and predicted numbers of death within 5 years and 10 years of diagnosis were calculated as mortality ratio (MR) with 95% CI constructed by exact

procedure [312]. MR was also calculated for different subgroups by country, period of diagnosis, age, race, and other clinical characteristics. Actual 5-year and 10-year survival rates were estimated by Kaplan-Meier analysis (observed survival), and compared with median CancerMath predicted survival. A difference of less than 3% would be considered reliable enough for clinical use as 10-year survival benefit of 3-5% is an indication for adjuvant chemotherapy [207]. The relationship of the predicted and observed 5-year and 10-year survival was also demonstrated by the calibration plot for the outcome and therapy calculator. Discriminative ability of predicted 5-year survival and 10-year survival from outcome and therapy calculator was evaluated by AUC using dataset with minimum 5-year and 10-year follow-up accordingly. Outcome calculator was further evaluated by C-statistics for the entire dataset regardless of follow-up time.

Conditional survival calculator

For patients who survived two years after diagnosis, predicted 5-year survival was compared with observed 5-year survival. Similarly predicted 10-year survival was compared with observed 10-year survival for patients who survived 5 years and 7 years respectively. Discriminative ability was evaluated by AUC.

7.3 Results

In total, 7064 female breast cancer patients were included. Tables 7.1-7.4 present clinical characteristics of 6807 patients with nodal status, 4517 patients with minimum 5-year follow-up, 1649 patients with 10-year follow-up, and 1538 patients with detailed treatment data and minimum of 5-years follow-up, respectively.

Nodal status calculator

A total of 6807 patients with nodal status data were selected for validation of nodal status calculator. In this dataset, 43.6% patients (n=2970) had at least one positive lymph node and the median predicted probability was 40.6%. In fact, cancerMath underestimated the probability of positive node for most of the subgroups (Table 7.1). The calibration plot (Figure 7.2) also illustrated underestimation except for the last two deciles of predicted probability. The discriminative ability of this calculator was fair, with overall AUC of 0.71 (95% CI, 0.70-0.72).

Table 7.1 Observed number of patients with positive lymph nodes and predicted probability of positive nodes among breast cancer patients.

	Number of patients with positive lymph nodes (percentage)	Predicted probability of positive nodes (median)
Overall	2970 (43.6%)	40.6%
Ethnicity		
Chinese	2062 (41.0%)	39.2%
Malay	511 (53.1%)	46.0%
Indian	312 (47.9%)	44.7%
Other	85 (51.8%)	39.5%
Country		
Malaysia	1460 (44.6%)	43.0%
Singapore	1510 (42.7%)	38.5%
Period of diagnosis		
1990-1994	58 (46.8%)	52.0%
1995-1999	258 (47.2%)	41.9%
2000-2003	755 (43.3%)	41.4%
2004-2007	1899 (43.2%)	39.8%
Age at diagnosis		
0-39	310 (46.3%)	47.1%
40-49	910 (44.6%)	42.9%
50-59	934 (43.5%)	41.4%
60-69	546 (42.0%)	36.7%
70+	270 (41.4%)	34.3%
Tumor size (mm)		
0-20	822 (28.1%)	26.4%
21-50	1678 (51.7%)	49.3%
51+	470 (74.1%)	79.2%
ER status		
Negative	1037 (44.8%)	43.5%
Positive	1854 (43.6%)	38.5%
Unknown	79 (33.3%)	44.5%
PR status		
Negative	1195 (45.0%)	42.1%
Positive	1511 (43.1%)	38.5%
Unknown	264 (41.0%)	44.2%
HER2 status		
Negative	1197 (41.7%)	39.2%
Equivocal	182 (42.4%)	39.2%
Positive	662 (50.3%)	45.0%
Unknown	929 (42.4%)	39.6%
Histology		
Ductal	2681 (45.1%)	41.5%
Lobular	150 (52.3%)	37.9%
Mucinous	34 (15.5%)	10.7%
Others	102 (29.0%)	35.8%
Unknown	3 (75.0%)	25.1%
Grade		
1	204 (24.0%)	21.8%
2	1278 (45.1%)	40.6%
3	1275 (51.8%)	46.4%
Unknown	213 (32.3%)	35.9%

Figure 7.1 Histogram of CancerMath predicted probability of positive nodes among 6807 patients with nodal status

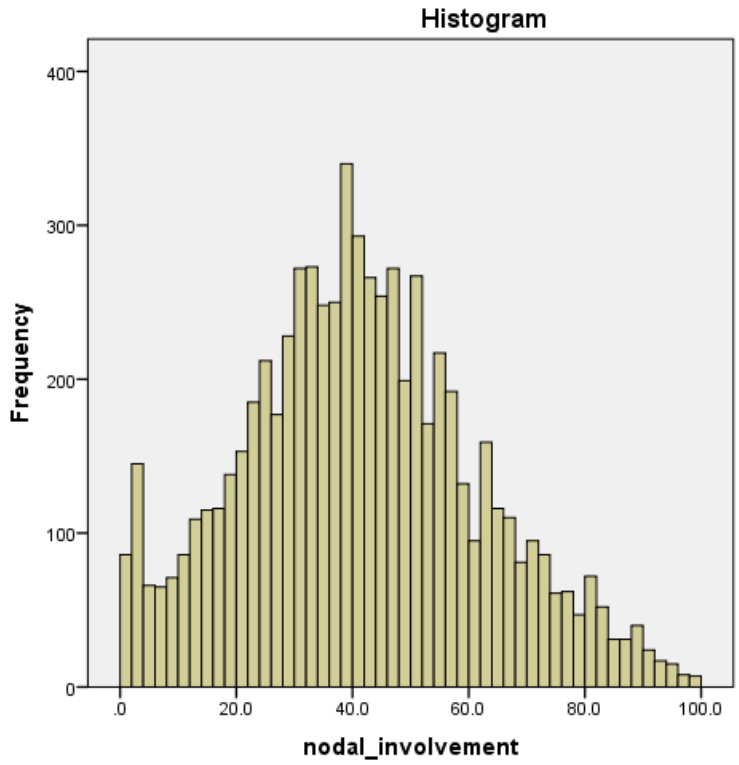
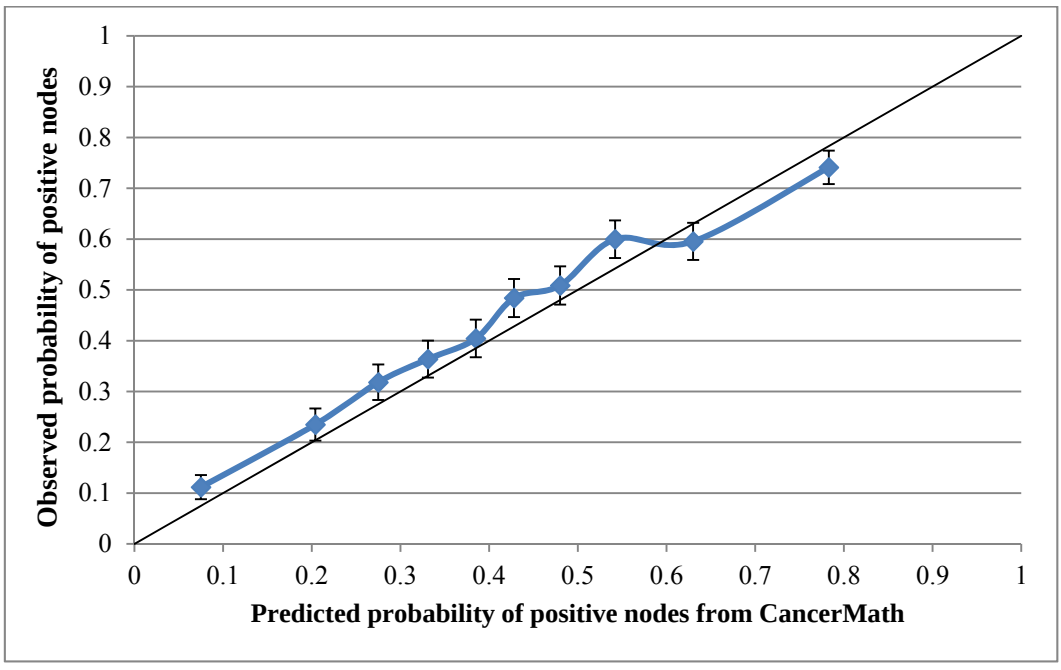


Figure 7.2 Calibration plot of observed probability of positive nodes with 95% confidence interval against predicted probability of positive nodes (median) by deciles of the predicted value



Outcome calculator

The observed number of deaths within 5 years after diagnosis was significantly higher than the predicted number of deaths (752 vs 667, MR=1.13, 95% CI 1.05-1.21). The number of observed and predicted number of deaths within 10 years after diagnosis was not significant (488 vs 454, MR=1.07, 95% CI 0.98-1.17). The absolute differences of 5-year and 10-year predicted and observed survival probabilities were 3.9% and 4.9%. Overestimation was more pronounced in Malaysian patients than in Singaporean patients (5.8% vs 2.5% for 5-year survival, and 8.0% vs 0.0% for 10-year survival). We also observed notable differences for cases diagnosed in earlier period and of younger age (Table 7.2 and 7.3). In addition, CancerMath significantly overpredicted survival for patients with unfavorable prognostic characteristics such as large tumor size, more positive nodes and ER negative tumor. For those with relatively better predicted survival, CancerMath predictions were similar to observed outcome (Figure 7.5 and 7.6). For example, the difference between 5-year predicted and observed survival was 17%, 3% and 1% for the first, fifth, and tenth deciles respectively. The AUC for 5-year and 10-year overall survival were 0.77 (95% CI, 0.75-0.79) and 0.74 (95% CI，0.71-0.76), respectively whereas the C-statistics was 0.74 (95% CI, 0.72-0.75). Both measures demonstrated fair discriminative ability.

Table 7.2 Observed and predicted 5-year overall survival from outcome calculator, stratified by patients' characteristics

	N	Observed deaths in 5 years	Predicted deaths in 5 years	Mortality Ratio (95% CI)	Observed 5-year survival (%) (std err)	Predicted 5-year survival (median) (%)	Absolute difference (%) (95% CI)
Overall	4517	752	667	**1.13(1.05,1.21)**	83.4 (0.006)	87.3	**3.9 (2.7,5.1)**
Ethnicity							
Chinese	3340	488	478	1.02(0.93,1.12)	85.4 (0.006)	88.0	**2.6 (1.4,3.8)**
Malay	654	143	104	**1.38(1.16,1.62)**	78.1 (0.016)	85.8	**7.7 (4.6,10.8)**
Indian	430	109	71	**1.54(1.26,1.85)**	74.7 (0.021)	85.1	**10.4 (6.3,14.5)**
Other	93	12	14	0.86(0.44,1.50)	87.1 (0.035)	87.3	0.2 (-6.7,7.1)
Country							
Malaysia	2143	423	331	**1.28(1.16,1.41)**	80.3 (0.009)	86.1	**5.8(4.0,7.6)**
Singapore	2374	329	336	0.98(0.88,1.09)	86.1 (0.007)	88.6	**2.5(1.1,3.9)**
Period of diagnosis							
1990-1994	140	41	22	**1.86(1.34,2.53)**	70.7 (0.038)	85.9	**15.2 (7.8,22.6)**
1995-1999	564	116	75	**1.55(1.28,1.86)**	79.8 (0.017)	87.9	**8.1 (4.8,11.4)**
2000-2003	1800	279	261	1.07(0.95,1.20)	84.5 (0.009)	87.8	**3.3 (1.5,5.1)**
2004-2007	2013	316	309	1.02(0.91,1.14)	84.3 (0.008)	87.2	**2.9 (1.3,4.5)**
Age at diagnosis							
0-39	493	101	64	**1.58(1.29,1.92)**	79.5 (0.018)	88.8	**9.3(5.8,12.8)**
40-49	1430	172	163	1.06(0.90,1.23)	88.0 (0.009)	90.6	**2.6(0.8,4.4)**
50-59	1412	224	194	**1.15(1.01,1.32)**	84.1 (0.010)	88.2	**4.1(2.1,6.1)**
60-69	776	126	130	0.97(0.81,1.15)	83.8 (0.013)	85.1	1.3(-1.2,3.8)

70+	406	129	117	1.10(0.92,1.31)	68.2 (0.023)	73.9	**5.7 (1.2,10.2)**
Tumor size (mm)							
0-20	1889	151	173	0.87(0.74,1.02)	92.0 (0.006)	92.9	0.9(-0.3,2.1)
21-50	2180	438	374	**1.17(1.06,1.29)**	79.9 (0.009)	84.8	**4.9 (3.1,6.7)**
51+	448	163	121	**1.35(1.15,1.57)**	63.6 (0.023)	73.6	**10.0(5.5,14.5)**
# of positive nodes							
0	2408	196	238	0.82(0.71,0.95)	91.9 (0.006)	91.7	-0.2(-1.4,1.0)
1-3	1068	195	165	**1.18(1.02,1.36)**	81.7 (0.012)	85.9	**4.2(1.8,6.6)**
4-9	533	159	122	**1.30(1.11,1.52)**	70.2 (0.020)	78.0	**7.8(3.9,11.7)**
10+	354	170	116	**1.47(1.25,1.70)**	52.0 (0.027)	67.4	**15.4(10.1,20.7)**
unknown	154	32	27	1.19(0.81,1.67)	79.2 (0.033)	86.6	**7.4(0.9,13.9)**
ER status							
Negative	1595	392	268	**1.46(1.32,1.61)**	75.4 (0.011)	85.2	**9.8 (7.6,12.0)**
Positive	2668	309	367	**0.84(0.75,0.94)**	88.4 (0.006)	88.8	0.4(-0.8,1.6)
Unknown	254	51	33	**1.55(1.15,2.03)**	79.9 (0.025)	88.6	**8.7(3.8,13.6)**
PR status							
Negative	1674	382	289	**1.32(1.19,1.46)**	77.2 (0.010)	84.8	**7.6(5.6,9.6)**
Positive	2174	241	285	**0.85(0.74,0.96)**	88.9 (0.007)	89.5	0.6(-0.8,2.0)
Unknown	669	129	93	**1.39(1.16,1.65)**	80.7 (0.015)	87.4	**6.7 (3.8,9.6)**
HER2 status							
Negative	1483	208	210	0.99(0.86,1.13)	86.0 (0.009)	88.0	**2.0(0.2,3.8)**
Equivocal	118	19	19	1.00(0.60,1.56)	83.9 (0.034)	87.4	3.5(-3.2,10.2)
Positive	790	172	147	**1.17(1.00,1.36)**	78.2 (0.015)	83.0	**4.8(1.9,7.7)**
Unknown	2126	353	292	**1.21(1.09,1.34)**	83.4 (0.008)	88.7	**5.3(3.7,6.9)**
Histology							
Ductal	3951	696	597	**1.17(1.08,1.26)**	82.4 (0.006)	87.0	**4.6 (3.4,5.8)**
Lobular	180	17	26	0.65(0.38,1.05)	90.6 (0.022)	87.5	-3.1(-7.4,1.2)

Mucinous	156	10	14	0.71(0.34,1.31)	93.6 (0.020)	94.7	1.1(-2.8,5.0)
Others	227	29	30	0.97(0.65,1.39)	87.2 (0.022)	89.5	2.3(-2.0,6.6)
Unknown	3	0	0	-	100	86.8	-13.2
Grade							
1	552	20	44	**0.45(0.28,0.70)**	96.4 (0.008)	94.7	**-1.7(-3.3,-0.1)**
2	1882	261	265	0.98(0.87,1.11)	86.1 (0.008)	88.2	**2.1(0.5,3.7)**
3	1591	402	288	**1.40(1.26,1.54)**	74.7 (0.011)	84.3	**9.6(7.4,11.8)**
Unknown	492	69	70	0.99(0.77,1.25)	86.0 (0.016)	87.4	1.4(-1.7,4.5)

Table 7.3 Observed and predicted 10-year overall survival from outcome calculator, stratified by patients' characteristics

	N	Observed deaths in 10 years	Predicted deaths in 10 years	Mortality Ratio (95% CI)	Observed 10-year survival (%)(std err)	Predicted 10-year survival (median) (%)	Absolute difference (%) (95% CI)
Overall	1649	488	454	1.07(0.98,1.17)	70.4 (0.011)	75.3	**4.9 (2.7,7.1)**
Ethnicity							
Chinese	1201	318	318	1.00(0.89,1.12)	73.5 (0.013)	76.8	**3.3 (0.8,5.8)**
Malay	251	100	74	**1.35(1.10,1.64)**	60.2 (0.031)	72.3	**12.1(6.0,18.2)**
Indian	174	64	55	1.16(0.90,1.49)	63.2 (0.037)	69.9	6.7 (-0.6,14.0)
Other	23	6	7	0.86(0.31,1.87)	73.9 (0.092)	77.1	3.2 (-14.8,21.2)
Country							
Malaysia	983	341	284	**1.20(1.08,1.34)**	65.3 (0.015)	73.3	**8.0 (5.1,10.9)**
Singapore	666	147	170	0.86(0.73,1.02)	77.9 (0.016)	77.9	0.0 (-3.1,3.1)
Period of diagnosis							
1990-1994	140	56	42	**1.33(1.01,1.73)**	60.0 (0.041)	72.5	**12.5(4.5,20.5)**
1995-1999	564	187	148	**1.26(1.09,1.46)**	66.8 (0.020)	76.0	**9.2 (5.3,13.1)**
2000-2002	945	245	264	0.93(0.82,1.05)	74.1 (0.014)	75.9	1.8 (-0.9,4.5)
Age at diagnosis							
0-39	232	82	58	**1.41(1.12,1.75)**	64.7 (0.031)	77.3	**12.6 (6.5,18.7)**
40-49	576	137	130	1.05(0.88,1.25)	76.2 (0.018)	80.2	**4.0 (0.5,7.5)**
50-59	493	141	129	1.09(0.92,1.29)	71.4 (0.020)	76.4	**5.0 (1.1,8.9)**
60-69	254	78	86	0.91(0.72,1.13)	69.3 (0.029)	68.4	-0.9 (-6.6,4.8)
70+	94	50	50	1.00(0.74,1.32)	46.8 (0.051)	50.1	3.3 (-6.7,13.3)

Tumor size (mm)							
0-20							
21-50	653	118	109	1.08(0.90,1.30)	81.9 (0.015)	86.8	**4.9 (2.0,7.8)**
51+	831	283	262	1.08(0.96,1.21)	65.9 (0.016)	70.6	**4.7 (1.6,7.8)**
	165	87	82	1.06(0.85,1.31)	47.3 (0.039)	50.6	3.3 (-4.3,10.9)
# of positive nodes							
0	867	147	161	0.91(0.77,1.07)	83.0 (0.013)	84.0	1.0 (-1.5,3.5)
1-3	407	143	120	**1.19(1.00,1.40)**	64.9 (0.024)	72.1	**7.2 (2.5,11.9)**
4-9	215	112	93	1.20(0.99,1.45)	47.9 (0.034)	58.2	**10.3 (3.6,17.0)**
10+	104	71	62	1.15(0.89,1.44)	31.7 (0.046)	39.9	8.2 (-0.8,17.2)
unknown	56	15	17	0.88(0.49,1.46)	73.2 (0.059)	73.5	0.3 (-11.3,11.9)
ER status							
Negative	637	224	197	1.14(0.99,1.30)	64.8 (0.019)	71.5	**6.7 (3.0,10.4)**
Positive	816	205	206	1.00(0.86,1.14)	74.9 (0.015)	78.2	**3.3 (0.4,6.2)**
Unknown	196	59	51	1.16(0.88,1.49)	69.9 (0.033)	76.8	**6.9 (0.4,13.4)**
PR status							
Negative	485	160	153	1.05(0.89,1.22)	67.0 (0.021)	70.7	3.7(-0.4,7.8)
Positive	564	128	136	0.94(0.79,1.12)	77.3 (0.018)	79.9	2.6 (-0.9,6.1)
Unknown	600	200	165	**1.21(1.05,1.39)**	66.7 (0.019)	74.1	**7.4 (3.7,11.1)**
HER2 status							
Negative	269	72	66	1.09(0.85,1.37)	73.2 (0.027)	78.3	5.1(-0.2,10.4)
Equivocal	13	6	4	1.50(0.55,3.26)	53.8 (0.138)	65.5	11.7 (-15.3,38.7)
Positive	335	113	110	1.03(0.85,1.24)	66.3 (0.026)	69.1	2.8 (-2.3,7.9)
Unknown	1032	297	273	1.09(0.97,1.22)	71.2 (0.014)	76.8	**5.6 (2.9,8.3)**
Histology							
Ductal	1418	445	401	**1.11(1.01,1.22)**	68.6 (0.012)	74.4	**5.8 (3.4,8.2)**
Lobular	78	18	21	0.86(0.51,1.35)	76.9 (0.048)	75.7	-1.2 (-10.6,8.2)

Mucinous	59	9	9	1.00(0.46,1.90)	84.7 (0.047)	91.2	6.5 (-2.7,15.7)
Others	91	16	22	0.73(0.42,1.18)	82.4 (0.040)	77.7	-4.7 (-12.5,3.1)
Unknown	3	0	1		100	74.4	-25.6
Grade							
1	200	22	31	0.71(0.44,1.07)	89.0 (0.022)	89.3	0.3 (-4.0,4.6)
2	668	188	176	1.07(0.92,1.23)	71.9 (0.017)	77.1	**5.2 (1.9, 8.5)**
3	510	196	172	1.14(0.99,1.31)	61.6 (0.022)	70.0	**8.4 (4.1,12.7)**
Unknown	271	82	76	1.08(0.86,1.34)	69.7 (0.028)	73.3	3.6 (-1.9,9.1)

Figure 7.3 Histogram of CancerMath predicted 5-year survival among 4517 patients with minimum 5-year follow up

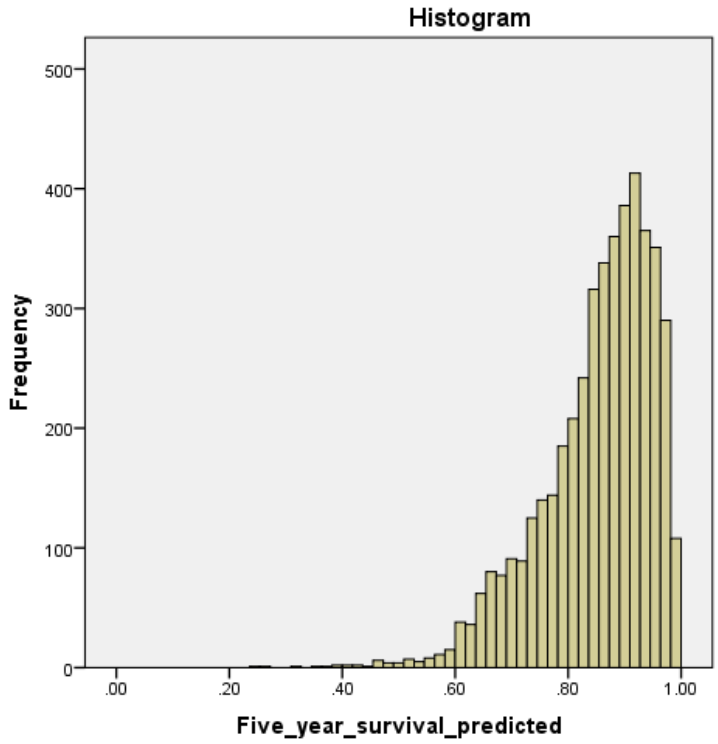


Figure 7.4 Histogram of CancerMath predicted 10-year survival among 1649 patients with minimum 10-year follow up

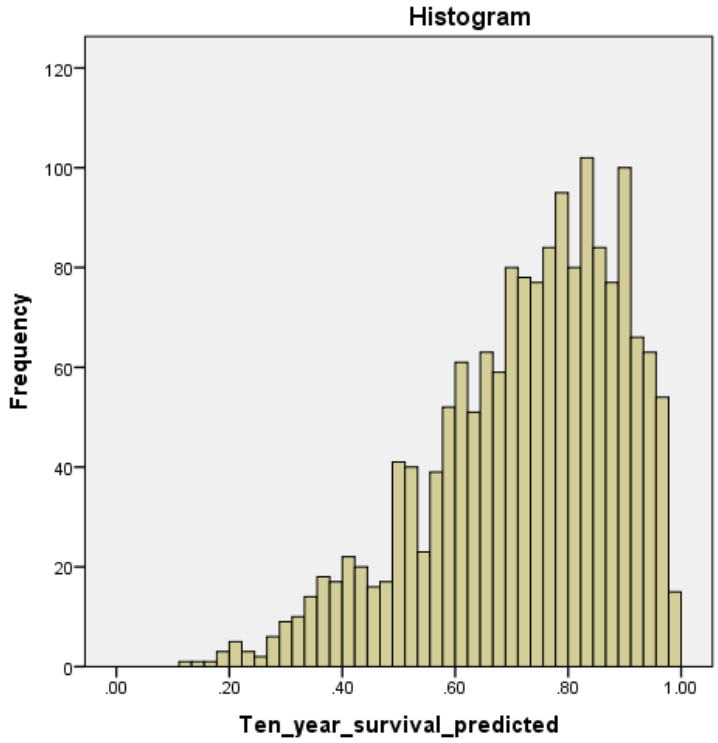


Figure 7.5 Calibration plot of observed survival with 95% confidence interval against predicted survival (median) by deciles of the predicted value for 5-year survival from outcome calculator

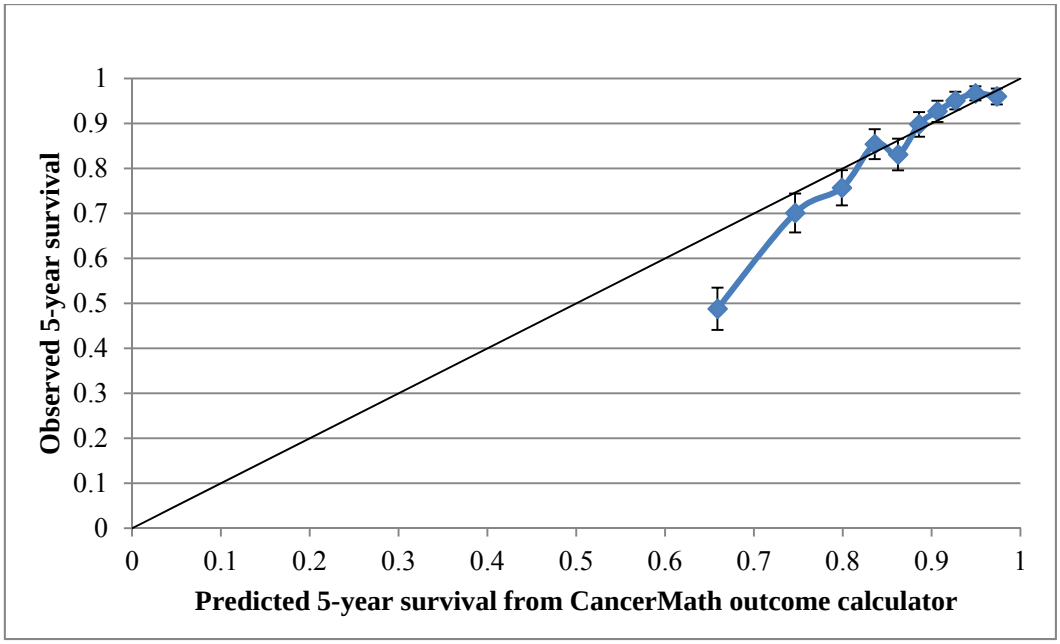
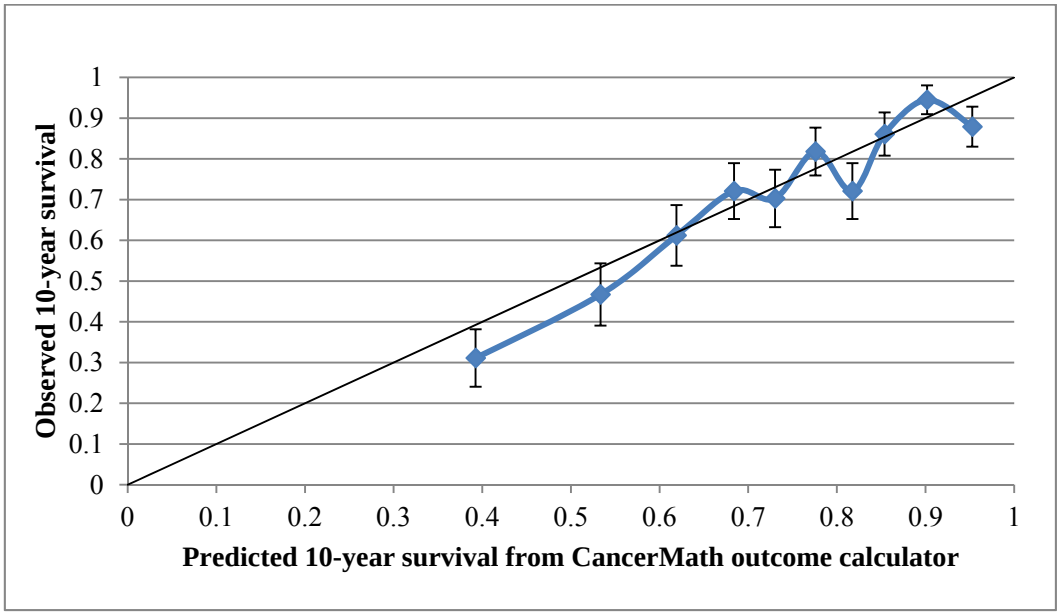


Figure 7.6 Calibration plot of observed survival with 95% confidence interval against predicted survival (median) by deciles of the predicted value for 10-year survival from outcome calculator



Therapy calculator

For the therapy calculator which was only validated in Malaysian patients, predicted survival was significantly higher than the observed survival for almost all subgroups, except for those diagnosed recently and with more favourable tumor characteristics (Table 7.4, Figure 7.7). The calculator showed fair discrimination at 5-year overall survival (AUC=0.73, 95% CI 0.70-0.77).

Conditional survival calculator

For patients who have survived 2 years since diagnosis, the predicted 5-year survival was 91.0% versus the observed survival of 88.3%. The AUC was 0.75 (95%CI, 0.73-0.77). For patients who have survived 5 years and 7 years, the predicted probability of surviving up to 10 years was 86.6% and 91.7% respectively. And the observed survival was 85.3% and 91.0% respectively. The AUC was 0.66 (95% CI, 0.62-0.70) and 0.63 (95% CI, 0.57-0.68) for 10-year survival.

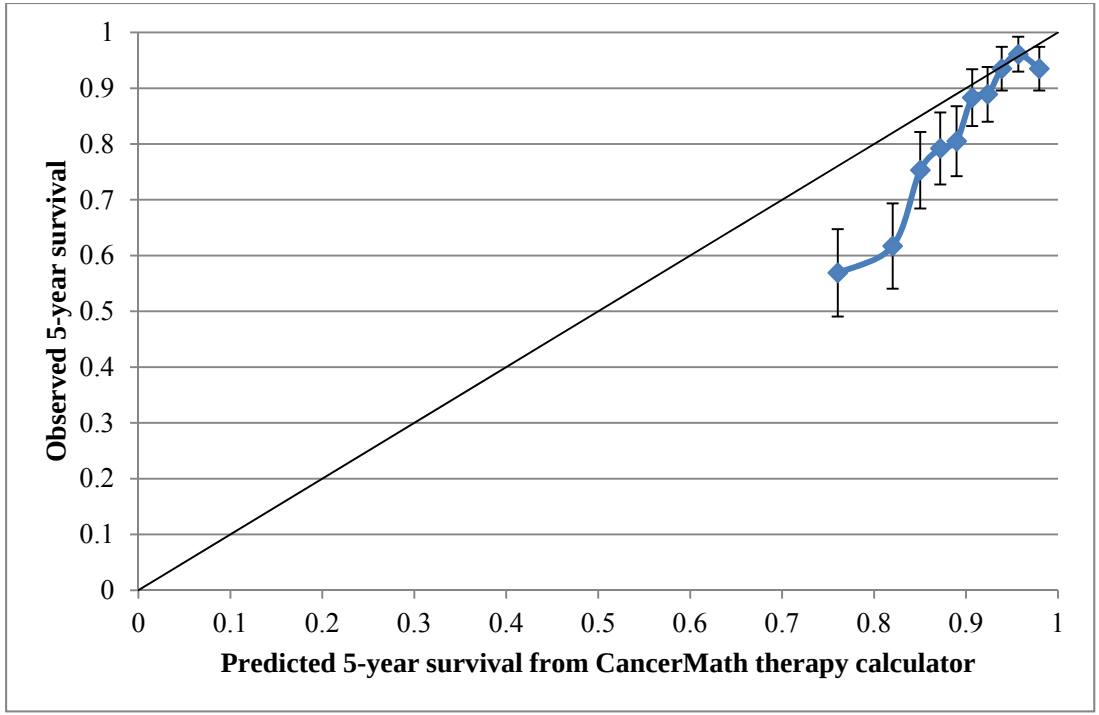
Table 7.4 Observed and predicted 5-year overall survival from therapy calculator, stratified by patients' characteristics

	N	Observed death in 5 years	Predicted death in 5 years	Mortality Ratio (95% CI)	Observed 5-year survival (%)(std err)	Predicted 5-year survival (median) (%)	Absolute difference (%) (95% CI)
Overall	1538	286	173	1.65(1.47,1.86)	81.4 (0.010)	89.8	**8.4(6.4,10.4)**
Ethnicity							
Chinese	1052	167	113	1.48(1.26,1.72)	84.1 (0.011)	90.4	**6.3(4.1,8.5)**
Malay	264	62	30	2.07(1.58,2.65)	76.5 (0.026)	89.4	**12.9(7.8,18.0)**
Indian	212	54	29	1.86(1.40,2.43)	74.5 (0.030)	87.2	**12.7(6.8,18.6)**
Other	10	3	1	3.00(0.62,8.77)	70.0 (0.145)	88.2	18.2(-10.2,46.6)
Period of diagnosis							
1990-1994	95	39	14	2.79(1.98,3.81)	58.9 (0.05)	86.8	**27.9 (18.1,37.7)**
1995-1999	374	93	40	2.33(1.88,2.85)	75.1 (0.022)	90.9	**15.8 (11.5,20.1)**
2000-2003	568	91	63	1.44(1.16,1.77)	84.0 (0.015)	89.7	**5.7 (2.8,8.6)**
2004-2007	501	63	56	1.13(0.86,1.44)	87.4 (0.015)	90.2	2.8 (-0.1,5.7)
Age at diagnosis							
0-39	205	55	17	3.24(2.44,4.21)	73.2 (0.031)	92.6	**19.4(13.3,25.5)**
40-49	515	74	41	1.80(1.42,2.27)	85.6 (0.015)	92.9	**7.3 (4.4,10.2)**
50-59	449	86	50	1.72(1.38,2.12)	80.8 (0.019)	89.4	**8.6 (4.9,12.3)**
60-69	271	43	40	1.08(0.78,1.45)	84.1 (0.022)	86.1	2.0 (-2.3,6.3)
70+	98	28	24	1.17(0.78,1.69)	71.4 (0.046)	77.4	6.0 (-3.0,15.0)
Tumor size (mm)							
0-20	547	51	39	1.31(0.97,1.72)	90.7 (0.012)	94.2	**3.5 (1.1,5.9)**
21-50	813	170	102	1.67(1.43,1.94)	79.1 (0.014)	88.5	**9.4 (6.7,12.1)**
51+	178	65	32	2.03(1.57,2.59)	63.5 (0.036)	82.8	**19.3 (12.2,26.4)**

# of positive nodes							
0	806	72	70	1.03(0.80,1.30)	91.1 (0.010)	92.4	1.3(-0.7,3.3)
1-3	389	83	46	1.80(1.44,2.24)	78.7 (0.021)	89.4	**10.7(6.6,14.8)**
4-9	192	64	30	2.13(1.64,2.72)	66.7 (0.034)	85.8	**19.1(12.4,25.8)**
10+	123	61	23	2.65(2.03,3.41)	50.4 (0.045)	82.3	**31.9(23.1,40.7)**
Unknown	28	6	4	1.50(0.55,3.26)	78.6 (0.078)	90.6	12.0 (-3.3,27.3)
ER status							
Negative	528	146	73	2.00(1.69,2.35)	72.3 (0.019)	87.2	**14.9(11.2,18.6)**
Positive	850	99	82	1.21(0.98,1.47)	88.4 (0.011)	91.7	**3.3 (1.1,5.5)**
Unknown	160	41	18	2.28(1.63,3.09)	74.4 (0.035)	89.8	**15.4 (8.5,22.3)**
PR status							
Negative	423	106	57	1.86(1.52,2.25)	74.9 (0.021)	87.4	**12.5(8.4,16.6)**
Positive	586	73	58	1.26(0.99,1.58)	87.5 (0.014)	91.6	**4.1 (1.4,6.8)**
Unknown	529	107	58	1.84(1.51,2.23)	79.8 (0.017)	90.2	**10.4 (7.1,13.7)**
HER2 status							
Negative	665	78	68	1.15(0.91,1.43)	88.3 (0.012)	91.1	**2.8 (0.4,5.2)**
Equivocal	35	7	4	1.75(0.70,3.61)	80.0 (0.068)	89.9	9.9 (-3.4,23.2)
Positive	418	84	53	1.58(1.26,1.96)	79.9 (0.020)	87.9	**8.0 (4.1,11.9)**
Unknown	420	117	48	2.44(2.02,2.92)	72.1 (0.022)	89.7	**17.6 (13.3,21.9)**
Histology							
Ductal	1346	270	155	1.74(1.54,1.96)	79.9 (0.011)	89.6	**9.7 (7.5,11.9)**
Lobular	71	7	7	1.00(0.40,2.06)	90.1 (0.035)	91.0	0.9 (-6.0,7.8)
Mucinous	58	1	4	0.25(0.01,1.39)	98.3 (0.017)	96.0	-2.3 (-5.6,1.0)
Others	63	8	7	1.14(0.49,2.25)	88.9 (0.040)	89.7	0.8 (-7.0,8.6)
Grade							
1	161	8	11	0.73(0.31,1.43)	95.0 (0.017)	95.6	0.6 (-2.7,3.9)
2	661	111	71	1.56(1.29,1.88)	83.2 (0.015)	90.5	**7.3 (4.4,10.2)**
3	433	119	59	2.02(1.67,2.41)	72.5 (0.021)	87.7	**15.2 (11.1,19.3)**
Unknown	283	48	32	1.50(1.11,1.99)	83.0 (0.022)	89.8	**6.8 (2.5,11.1)**

Chemotherapy							
No chemotherapy	440	58	53	1.09(0.83,1.41)	86.8 (0.016)	90.4	**3.6 (0.5,6.7)**
1st Gen	162	49	21	2.33(1.73,3.08)	69.8 (0.036)	88.1	**18.3 (11.2,25.4)**
2nd Gen	915	174	97	1.79(1.54,2.08)	81.0 (0.013)	90.0	**9.0 (6.5,11.5)**
3rd Gen	21	5	2	2.50(0.81,5.83)	76.2 (0.093)	90.8	14.6 (-3.6,32.8)
Hormone-therapy							
No	398	108	51	2.12(1.74,2.56)	72.9 (0.022)	87.7	**14.8 (10.5,19.1)**
Yes	1140	178	122	1.46(1.25,1.69)	84.4 (0.011)	90.8	**6.4 (4.2,8.6)**

Figure 7.7 Calibration plot of observed survival with 95% confidence interval against predicted survival (median) by deciles of the predicted value for 5-year survival from therapy calculator



7.4 Discussion

Many prognostic tools have been developed over the past two decades to aid clinical decision making for breast cancer patients. This study validated four different prognostic calculators provided by CancerMath in the Singapore Malaysia Hospital Based Breast Cancer Registry. The discrimination was fair for nodal status calculator. CancerMath outcome, therapy and conditional survival calculator also moderately discriminated between survivors and nonsurvivors at 5 years and 10 years after diagnosis. It however consistently overestimated survival for this cohort of Southeast Asian patients, especially for those with poor predicted prognosis, as assessed by the calibration plot.

CancerMath was previously built and validated using SEER data and patients diagnosed at Massachusetts General and Brigham and Women's Hospitals [189]. It was shown to be highly accurate and the difference between observed and predicted survival was within 2% for 97% of the patients in the validation set [189]. Our study is the first one to independently validate CancerMath outside its initial study population and is also the largest validation study of a Western-derived breast cancer prognostic model in Asia. We demonstrated that CancerMath overpredicted survival by more than 3% for almost all clinical and pathological subgroups. The findings were similar to previous validation studies of Adjuvant! Online conducted in Asia. In the Malaysian, Korean, and Taiwanese studies, the predicted and observed 10-year overall survival differed by 6.7%, 11.1%, and 3.9% correspondingly [197-199]. The AUC was 0.73 (95% CI, 0.69- 0.77) in the Malaysian study and hence very close to the AUC of CancerMath reported in the present study [197]. Furthermore the prediction was too optimistic for young

patients in almost all validation studies of Adjuvant! Online [193, 196-198]. Although adjustment of 1.5-fold increase in risk was added to Adjuvant! Online version 7.0 for patients younger than 36 years and with ER positive breast cancer as stated in its help files, overprediction was still found in recent validation studies [193, 197, 198]. Our findings from current validation of CancerMath also suggest that correction for young age at diagnosis is needed.

The selection of patients for validation can partially explain the discrepancy in observed and predicted survival. CancerMath has only been validated among patients with tumor size no more than 50mm and positive nodes no more than seven [313]. In our validation dataset, 10% of patients had tumor size larger than 50mm and 8% had more than ten positive nodes. However even for patients with tumor size in between 20mm and 50mm and one to three positive nodes, the difference between the predicted and observed survival was more than 3%. In general, Asian patients are more likely to present with unfavorable prognostic features such as young age, negative hormone receptor status, HER2 overexpression, and more advanced stage compared to their western counterparts [102, 314, 315]. In our current analysis, reduced agreement was observed for patients with poorer outcome as illustrated by the calibration plot. CancerMath also performed poorly in Malaysian patients than Singaporean patients due to higher proportion of patients in advanced stages in Malaysia [316]. Such limitation of CancerMath may restrict its use to patients with better prognostic profile only. Furthermore CancerMath therapy calculator applies the same amount of risk reduction from adjuvant therapy as Adjuvant! Online, which was estimated from meta-analysis of clinical trials mainly conducted in the Western population [147, 188, 189, 310]. However non-adherence to treatment is more common among Asian women [317-322]. Studies

also reported different drug metabolism and toxicity induced by chemotherapy between Asian and Caucasian patients [323]. These evidences may imply CancerMath overestimate the effect of treatment in Asian patients.

Another possible explanation of suboptimal performance of CancerMath and also the limitation of our study would be missing data on ER (6%), PR (15%), HER2 status (47%), and tumor grade (11%). For patients with complete information on required predictors (N=1872), the predicted and observed 5-year survival was 86.0% and 82.5%. The difference was similar to what we observed in the entire dataset. Therefore the impact of missing data is relatively small on performance of CancerMath.

Several gene expression profiling assay, such as MammaPrint [105] and Oncotype Dx [106] are currently available in the market for breast cancer prognostication and treatment decision. However these tools do not incorporate clinical and histological factors which are readily available or relatively cheap to obtain. Due to the high cost of these tests and larger proportion of patients with high predicted risk in Asia [324, 325], the clinical utility is uncertain in this region. Therefore traditional prognostic model using clinicopathologic factors seems more reasonable in our local setting.

In conclusion, we found that the discriminative ability and calibration of CancerMath calculators was modest in Southeast Asian patients. Our results suggested that CancerMath was more suitable for patients diagnosed with favourable disease and received adequate treatment.

Chapter 8 Predicting survival of *de novo* metastatic breast cancer in Asian women: Systematic review and validation study of prognostic tools

Originally published by PLoS One. [Miao H, Hartman M, Bhoo-Pathy N, Lee SC, Taib NA, Tan EY, Chan P, Moons KG, Wong HS, Goh J, Rahim SM, Yip CH, Verkooijen HM: PLoS One 9(4), 2014 Apr 2.]

8.1 Motivation

Asian women are more likely to be diagnosed with late stage disease compared to their Western counterparts. Approximately 10% to 25% of Asian breast cancer patients present with *de novo* metastatic disease, compared to 3% to 5% in Europe and United States [315, 326-329]. In addition, metastatic lesions in Asian women are larger and often involve multiple sites [200].

Metastatic breast cancer is incurable. Median survival rates range from one to four years, but on an individual level, survival times of up to 15 years have been reported [330-337]. While recent studies suggest that surgical removal of primary breast tumor has a positive impact on the survival of *de novo* metastatic patients [145, 338, 339], systemic therapy, is the main treatment. Due to advances in locoregional and systemic treatment and due to the detection of small, solitary metastases, survival has improved over time, especially in patients with hormone receptor-positive tumors [334, 337].

Accurate assessment of individual prognosis of patients with *de novo* metastatic breast cancer is needed for treatment decision making. In addition, like all patients with cancer, women with distant metastases want to know their prognosis [340]. As clinicians are known to be overoptimistic in predicting survival [341], prediction

rules can be useful for this heterogeneous group of patients with different treatment options. Although many multivariable prognostic indices have been developed for breast cancer in the last two decades, the majority are not applicable to patients with *de novo* metastatic disease [187-189]. In this study, we aim to identify prediction tools which can be used for prognostication of patients with *de novo* metastatic breast cancer and externally validate their performance in the Singapore Malaysia Hospital Based Breast Cancer Registry.

8.2 Methods

Systematic review

Our first step was to perform a systematic review of the available literature, according to the PRISMA guidelines [342]. A free text search was performed on 13 August 2013 to identify eligible studies using MEDLINE and EMBASE electronic database. Our search strategy included search terms and synonyms for prognostic models and the following string was used: ((metastatic breast cancer) AND ((prognostic scor* OR prognostic index OR nomogram OR predictive model OR validation OR validate OR prognostic model OR predictor) AND (scor* OR index OR model OR predict* OR nomogram OR validat*))) NOT (expression profiling OR microarray* OR proteomic OR affymetrix). After reviewing the titles and abstracts, full text was selected applying predefined inclusion and exclusion criteria. Included were studies presenting multivariable models, with the aim to predict overall survival of metastatic breast cancer patients. We excluded animal models or clinical trials on treatment efficacy, as well as studies which used disease-free, progression-free survival or response to treatment as the only outcome of interest. Etiological studies which only assessed the effect size of one specific prognostic

factor or only evaluated the prognostic value of a single biomarker were not included. We also excluded prediction tools developed for patients with metastases from various primary cancers. Prognostic tools for patients with advanced cancer nearing the end of life or tools specific for recurrent metastatic breast cancer were not included as these patients have been exposed to multiple chemotherapy regimens and are often treatment resistant. Two studies which validated previously published models in metastatic breast cancer patients were excluded. Additional articles were retrieved by cross-referencing. Details regarding the author, year of publication, study design, model variables and performance measures were extracted if available. Quality of the selected publications was assessed using items listed in the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement, which were relevant to our study [343].

Validation set

Patients diagnosed with *de novo* metastatic breast cancer between 2000 and 2010 were retrieved from Singapore Malaysia Hospital Based Breast Cancer Registry. They were followed up from the date of diagnosis until the date of death or date of last contact whichever came first. The date of last contact was 1 November 2010 for UMMC patients, 1 July 2011 for NUH patients and 1 October 2012 for TTSH patients.

Statistical analysis

In the validation set, we investigated the pattern of missing data and assumed that data missingness was related to at least one other variable but not dependent on value of the observation itself, i.e. missing at random [344]. A total number of 230 (36%) individuals had complete data on all variables used in validation and 90 (14%)

cases had 3 or more variables missing. On average, each individual had 1.13 variables missing (standard deviation=1.22), ranging from 0 to 5. Missing values were imputed once using regression imputation [344].

For each individual patient, we calculated the prognostic score for the different prognostic models/indices except for those developed by recursive partitioning analysis [345] and artificial neural network [346], as terminal nodes were missing in our dataset or algorithm was not provided to allow calculation of prognostic scores. For models including performance status, a variable that was not captured in our database, we assumed all patients to be fit at the time of diagnosis, i.e. 0 on Zubrod scale, which is the same as the Eastern Cooperative Oncology Group (ECOG) and the WHO scale, and 100 on the Karnofsky performance status (KPS) scale. In order to check this assumption, we retrieved comorbidity data from the medical records of a subset of 87 NUH patients who diagnosed after 2006. We also assumed the best case scenario for lactate dehydrogenase (LDH). For brain metastasis models, a score of zero (best case scenario) was assigned to the largest brain metastasis dimension in Marko et al.'s model. We assumed no trastuzumab use for HER2 positive patients in Ahn et al.'s model, as in Singapore and Malaysia trastuzumab use was rare during the time of our study. Since our study population consisted of patients who were metastatic at presentation, disease free interval was set as zero for all women.

The distribution of each prognostic score was then divided into tertiles with the exception for Rabinovich's model, for which were only two possible combinations. We compared the survival of low, intermediate and high-risk score patients by plotting the Kaplan-Meier survival curves for each tertile. Median survival and 95% CIs were obtained for different groups and differences were tested by log-rank test

and log-rank test for trend. The discrimination ability of the models was assessed by C-statistic. For models with C-statistic larger than 0.6, 1-year, 2-year and 3-year cumulative survival probabilities were plotted for each quintile of the prognostic score.

8.3 Results

Systematic review

The search strategy resulted in 1298 titles (Figure 8.1). Forty-eight full text articles were selected after screening the titles and abstracts and two articles were added by cross-referencing. A total of 16 prognostic indices met our inclusion criteria. Eight models were developed for patients with metastatic breast cancer in general, seven for patients with brain metastasis from breast cancer and one for breast cancer patients with metastatic spinal cord compression [347-362]. All prognostic indices were designed for both *de novo* and recurrent metastatic breast cancer patients (Table 8.1). Study sizes ranged from 83 to 619 patients, with a median study size of 246 patients. The median survival from time of detection of metastasis ranged from 9.6 to 22 months. Cox regression incorporated time-to-event data and all-cause mortality as outcome was used for model development in 13 studies. Three studies conducted recursive partitioning analysis and one used artificial neural network. For Cox regression modeling, forward or backward stepwise selection with different cut-off P-values, either 0.05 or 0.1 was applied to identify final predictors.

Performance status, ER status, metastatic site(s) and disease free interval were the most common prognostic factors included in the different models. Performance status was measured on different scales, i.e. five studies used Zubrod/ECOG/WHO score while six models for brain metastasis used KPS [348, 350, 352, 354, 356-362]. Model coefficients or hazard ratios were presented in all Cox regression models. Six studies transformed the model into a scoring system for easy calculation of predicted survival and three studies developed a nomogram [347, 351, 352, 354, 356-359, 362]. Recursive decision tree was constructed from recursive partitioning analysis in two studies [363, 364]. Only five studies evaluated the discrimination of their models using C-statistic or AUC [350, 353, 354, 358, 359], which ranged from 0.67 to 0.74 (moderate discrimination). Calibration was assessed by plotting predicted versus observed survival for only two models, which turned out to be well calibrated [358, 359]. Four studies conducted internal validation using random subset of data, ten-fold cross-validation and bootstrapping with 200 and 1000 resamples [353, 358, 359, 362, 364]. Temporal validation of the model using data collected from the same hospital but later than those in the development set was conducted in four studies [348, 350, 352]. Five models were externally validated in other hospitals or outside the original country [351, 354, 358, 359, 363]. Quality of the selected publications is summarized in Table 8.2.

Figure 8.1 Flow chart of study selection process.

n = number of studies.

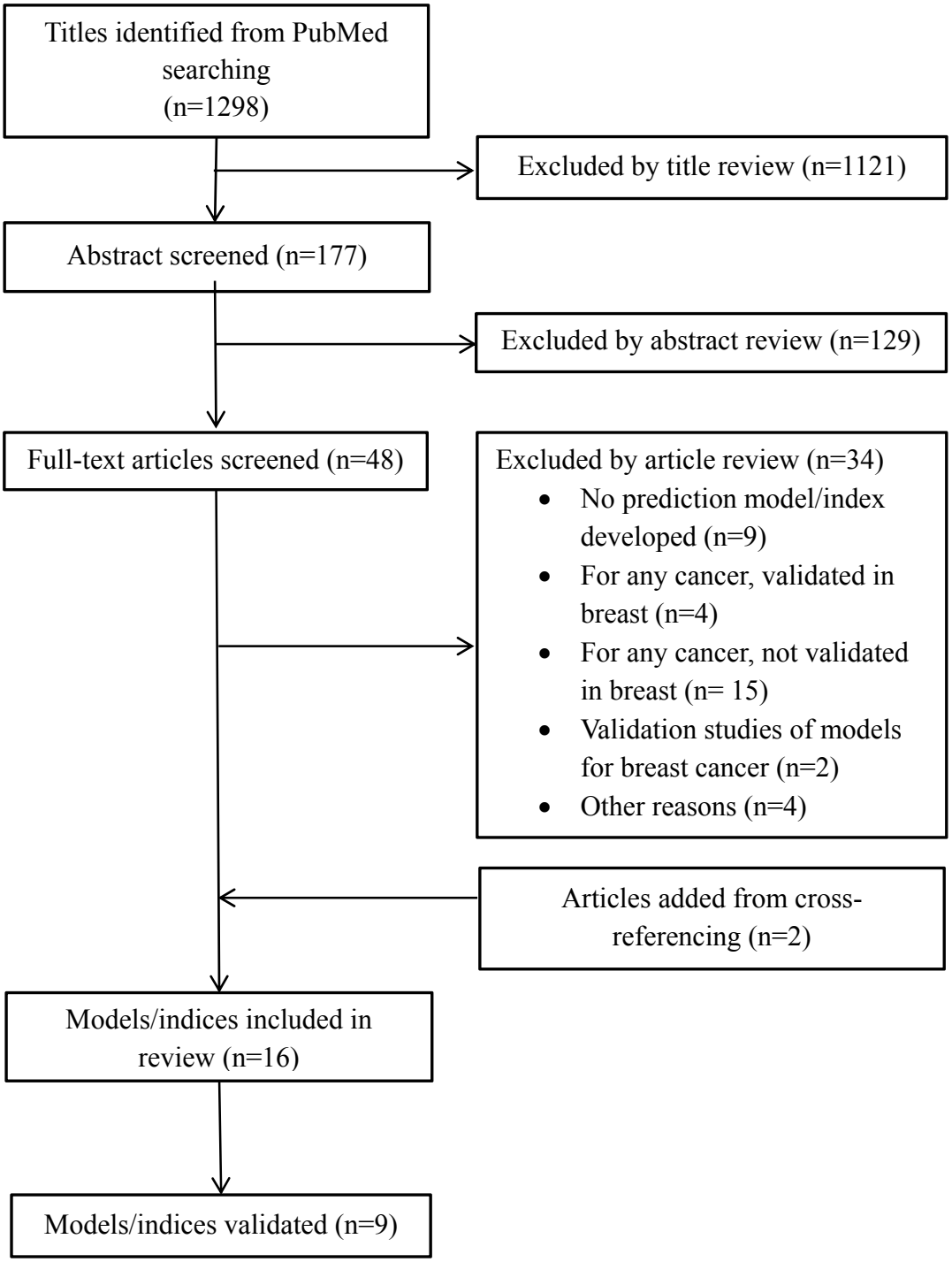


Table 8.1 Study characteristics of prognostic models for metastatic breast cancer patients

Authors	Year of publication	Number of patients	Country	Setting	Period of diagnosis	Median survival	Predictors	Analysis	Discrimination	Validation
Nash et al.	1980	138	USA	Single institution	1973-1977	17 months	age, number of metastatic site(s)	Cox Regression	Not reported	No
Hortobagyi et al.	1983	619	USA	Single institution	1973-1976	22 months	LDH, PS,site(s) of metastasis, radiotherapy, ALKP and extent of disease	Cox Regression	Not reported	Temporal
Williams et al.	1986	191	UK	Single institution, patients without brain metastasis	1974-1984	Not reported	Grade, ER status, DFI, site(s) of initial metastasis	Cox Regression	Not reported	External and temporal
Rabinovich et al.	1992	362	Argentina	Multiple institutions	1978-1985	21 months	PS, visceral involvement	Cox Regression	C-statistic = 0.72	Temporal
Yamamoto et al.	1998	233	Japan	Multiple institutions	Not available	21.5 months	adjuvant chemotherapy, presence of distant lymph nodes, liver metastasis, LDH and DFI	Cox Regression	Not reported	External
Ryberg et al.	2001	469	Denmark	single institution	1983-1992	14.7 months	Metastatic site(s), LDH, age, ER status and PS	Cox regression	Not reported	Temporal
Giordano et al.	2011	311	USA	Single institution	2004-2009	34.0, 28.3, 20.5 and 8.1 months for four risk	ER, PR, HER2 status, visceral metastasis, bone metastasis, number of metastatic site(s),	artificial neural network	C-statistic = 0.73	Internal

						groups based on CTC	therapy type, line of treatment; and CTC count			
Giordano et al.	2013	236	USA	Single institution	2002- 2009	Not reported	age, hormone receptor and HER2 status, visceral metastases, PS and CTC	Cox Regression	C-statistic = 0.74	External
Le Scodan et al.	2007	117	France	Single institution, patients with brain metastasis	1998-2003	5 months	RTOG RPA, Lymphocyte count, hormone receptor status	Cox Regression	Not reported	No
Nieder et al.	2009	83	Norway, Germany	2 institutions, patients with brain metastasis	2002-2007	16.0, 5.5 and 2.7 months for low, medium and high risk groups	KPS, extracranial metastases, multiple brain metastasis and DFI	Cox Regression	Not reported	No
Sperduto et al.	2012	400	USA	11 institutions, patients with brain metastasis	1993-2010	13.8 months	KPS, age, ER, PR and HER2 status	Cox regression, RPA	Not reported	External
Ahn et al.	2012	171	Korea	Single institution, patients with brain metastasis	2000–2008	9.6 months	KPS, extracranial metastases, age, trastuzumab, ER, PR and HER2 status	Cox Regression	Area under a curve=0.73	Internal and external
Marko et al.	2012	261	USA	Single institution, patients with brain metastasis	1999-2008	16.2 months	age, KPS, Non-CNS and number of CNS metastases, largest dimension brain metastasis,	Cox Regression	C-statistic = 0.67	Internal

							ER, PR, HER2, breast cancer stage			
Le Scodan et al.	2012	130	France	Single institution, patients with brain metastasis	1998-2006	7.43 months	KPS, age, trastuzumab,, ER,PR,HER status and lymphocyte count	RPA	Not reported	No
Niwińska et al.	2012	441	Poland	Single institution, patients with brain metastasis	2003-2009	7 months	KPS, number of brain metastases and extracranial metastasis	RPA	Not reported	No
Rades et al.	2013	255	Germany, Netherland , UK, Bosnia Herzegovi na	Multiple institutions , patients with metastatic spinal cord compressio n	1995-2011	Not reported	PS, ambulatory status, other bone metastases, visceral metastases, interval to radiotherapy, time of developing motor deficits	Cox Regression	Not reported	Internal

Abbreviation: LDH, Lactate dehydrogenase; PS, Performance status (Zubrod/ECOG/WHO score); ALKP, alkaline phosphatase; DFI, disease free interval; KPS, Karnofsky performance score; CNS, Central nervous system; ER, Estrogen receptor; PR, Progesterone receptor; HER2, Human epidermalgrowth factor receptor 2; CTC, circulating tumor cells; RTOG, Radiation Therapy Oncology Group; RPA, recursive partitioning analysis

Table 8.2 Summary of quality assessment of publications selected for validation (Y, yes (presented in study)

Authors	Inclusion and exclusion criteria clearly described	Outcome (survival) clearly described	Predictors clearly described	Loss of follow-up <20%	Characteristics of patients clearly described	Discrimination & calibration	Internal or external validation
Nash et al.	Y			Y	Y		
Hortobagyi et al.		Y	Y	Y	Y		Y
Williams et al.	Y	Y	Y				Y
Rabinovich et al.	Y	Y	Y	Y	Y	Y	Y
Yamamoto et al.		Y	Y	Y	Y		Y
Ryberg et al.			Y		Y		Y
Giordano et al. 2011	Y		Y		Y	Y	Y
Giordano et al. 2013	Y	Y			Y	Y	Y
Le Scodan et al. 2007	Y	Y	Y	Y	Y		
Nieder et al.	Y			Y	Y		
Sperduto et al.	Y	Y	Y	Y	Y		Y
Ahn et al.	Y	Y	Y		Y	Y	Y
Marko et al.	Y	Y	Y		Y	Y	Y
Le Scodan et al. 2012	Y	Y	Y		Y		
Niwińska et al.		Y	Y		Y		
Rades et al.	Y		Y		Y		Y

Validation

Our validation set included 642 Asian *de novo* metastatic breast cancer patients with a median age of 53 years (range, 24-94). Patient characteristics are reported in Table 8.3. Over a follow-up period of 1267.6 person-years, 492 patients had died and the median survival time was 19 months (95% CI, 16.5-21.5). The 1-year, 2-year and 3-year survival rates were 62%, 43% and 31% respectively. Half of the patients had more than one metastatic site involved and the majority did not receive any surgery or radiotherapy. Chemotherapy and hormone therapy were administered to 53% and 32% of the study population respectively. Among the 87 NUH patients with comorbidity data, hypertension (30%) and diabetes (23%) were the most common medical conditions. Less than 10% of this group was suffering from coronary heart disease (7%), stroke (2%), chronic obstructive pulmonary disease (3%) and renal failure (1%) and 6% of the patients have more than two comorbidities.

We validated all models that used Cox regression, with the exception of the models developed by Hortobagyi et al., Giordano et al., Le Scodan et al. and Rades et al. because the key predictors alkaline phosphatase (ALKP), circulating tumor cell (CTC), lymphocyte count and metastasis to spine were not available. Only Williams et al.'s, Yamamoto et al.'s, Rabinovich et al.'s and Ryberg et al.'s models were able to significantly discriminate between different risk groups in terms of overall survival based on log-rank test (Figure 8.2). The median survival for the low-risk group, intermediate-risk group and high-risk group classified according to Williams et al.'s model was 30 months, 21 months and 10 months respectively. For Rabinovich et al.'s model with two possible combinations, the median survival was 27 months and 16 months for the low and high risk groups. For Ryberg et al.'s

model, the median survival was 29, 17 and 10 months respectively for the three groups. However the log-rank for trend test was not significant for Yamamoto et al.'s model as the median survival was 17 months for the low risk group, 24 months for the medium risk group and 15 months for the high risk group.

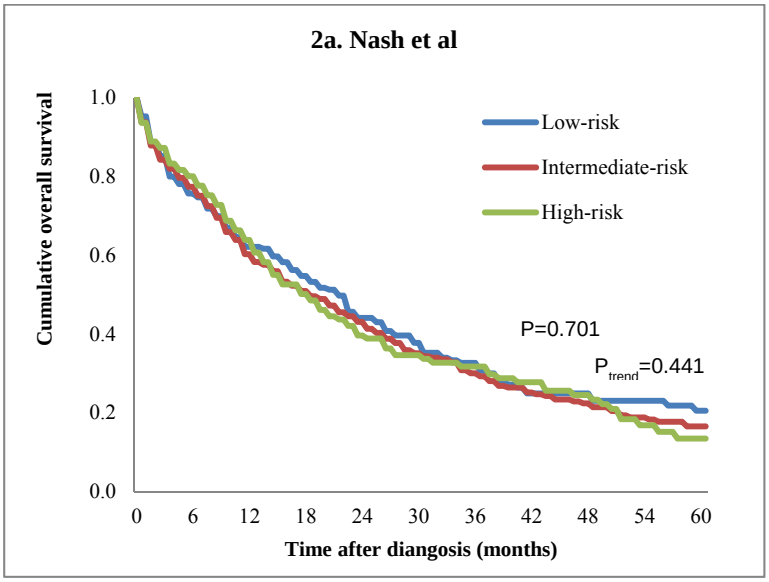
In our cohort, discrimination of the different models was poor to fair, with C-statistics ranging from 0.51 to 0.63 (Table 8.4). The model with the highest discriminatory ability was the model developed by Williams et al. (C-statistic=0.63, 95% CI 0.60-0.66), followed by Ryberg et al. (C-statistic=0.61, 95% CI 0.59-0.64). A notable decreasing trend of 1-year, 2-year and 3-year cumulative survival probabilities was observed for the five risk groups (quintiles, Figure 8.3). For Williams et al.'s model, the 3-year survival probabilities for the lowest and highest risk group were 49% (95% CI, 39%-58%) and 10% (95% CI, 4%-16%) respectively. For Ryberg et al.'s model, 3-year survival probabilities were 53% (95 % CI, 45%-61) and 13% (95 % CI, 7%-19%) for the low versus high risk groups respectively.

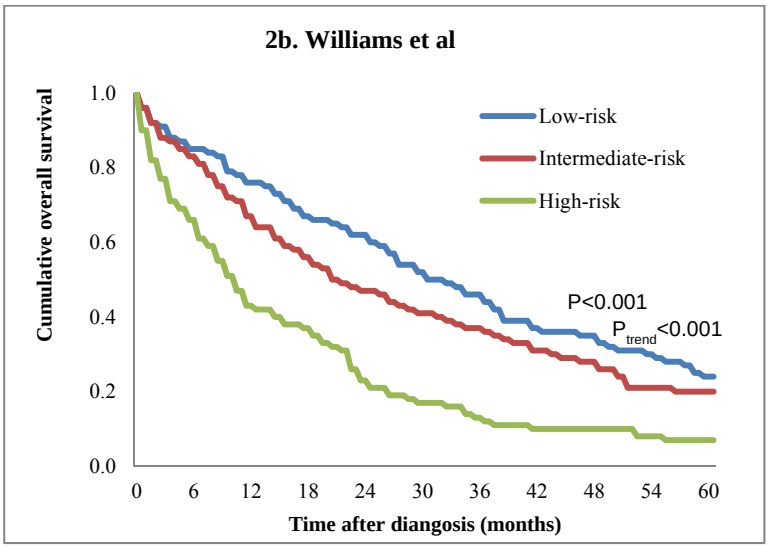
Table 8.3 Characteristics of *de novo* metastatic breast cancer patients identified at NUH, TTSH and UMMC, 2000-2010

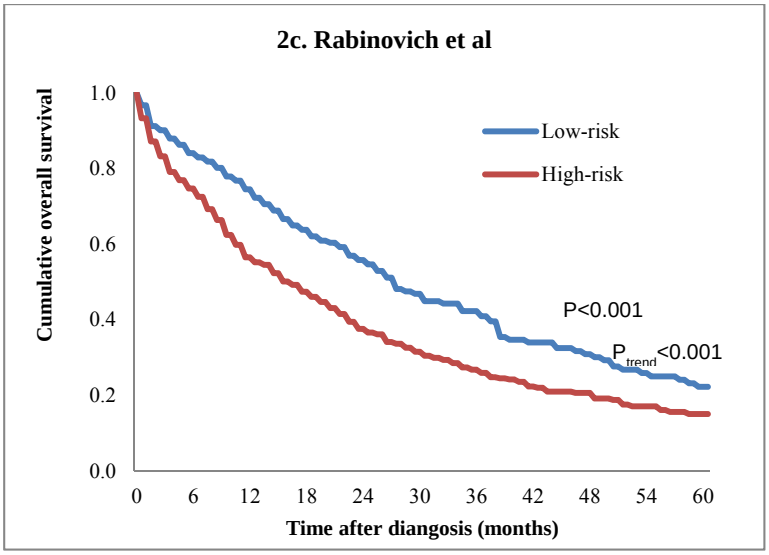
		UMMC	NUH	TTSH	Overall
Total		266 (41.4%)	156 (24.3%)	220 (34.3%)	642
Median Survival in months (95% CI)		14.0 (11.7-16.3)	28.0 (20.9-35.1)	18.0 (12.2-23.8)	19.0 (16.5-21.5)
Median age at diagnosis in years (range)		50 (24-83)	53 (28-80)	58 (30-94)	53 (24-94)
Median tumor size in mm (range)		100 (5-300)	40 (2-210)	60 (2-200)	60 (2-300)
Ethnicity	Chinese	148 (55.6%)	95 (60.9%)	152 (69.1%)	395 (61.5%)
	Malay	88 (33.1%)	38 (24.4%)	39 (17.7%)	165 (25.7%)
	Indian	30 (11.3%)	12 (7.7%)	15 (6.8%)	57 (8.9%)
	Others	0 (0.0%)	11 (7.1%)	14 (6.4%)	25 (3.9%)
Grade	1	2 (0.8%)	5 (3.2%)	3 (1.4%)	10 (1.6%)
	2	53 (19.9%)	64 (41.0%)	40 (18.2%)	157 (24.5%)
	3	63 (23.7%)	70 (44.9%)	41 (18.6%)	174 (27.1%)
	Unknown	148 (55.6%)	17 (10.9%)	136 (61.8%)	301 (46.9%)
ER status	Negative	102 (38.3%)	51 (32.7%)	81 (36.8%)	234 (36.4%)
	Positive	116 (43.6%)	103 (66.0%)	129 (58.6%)	348 (54.2%)
	Unknown	48 (18.0%)	2 (1.3%)	10 (4.5%)	60 (9.3%)
PR status	Negative	104 (39.1%)	62 (39.7%)	130 (59.1%)	296 (46.1%)
	Positive	63 (23.7%)	92 (59.0%)	80 (36.4%)	235 (36.6%)
	Unknown	99 (37.2%)	2 (1.3%)	10 (4.5%)	111 (17.3%)
HER2 status	Negative	64 (24.1%)	71 (45.5%)	75 (34.1%)	210 (32.7%)
	Positive	77 (28.9%)	24 (15.4%)	57 (25.9%)	158 (24.6%)
	Equivocal	20 (7.5%)	12 (7.7%)	17 (7.7%)	49 (7.6%)
	Unknown	105 (39.5%)	49 (31.4%)	71 (32.3%)	225 (35.0%)
Site(s) of metastases	Bone only	57 (21.4%)	25 (16.0%)	46 (20.9%)	128 (19.9%)
	Lung only	45	11	30	86

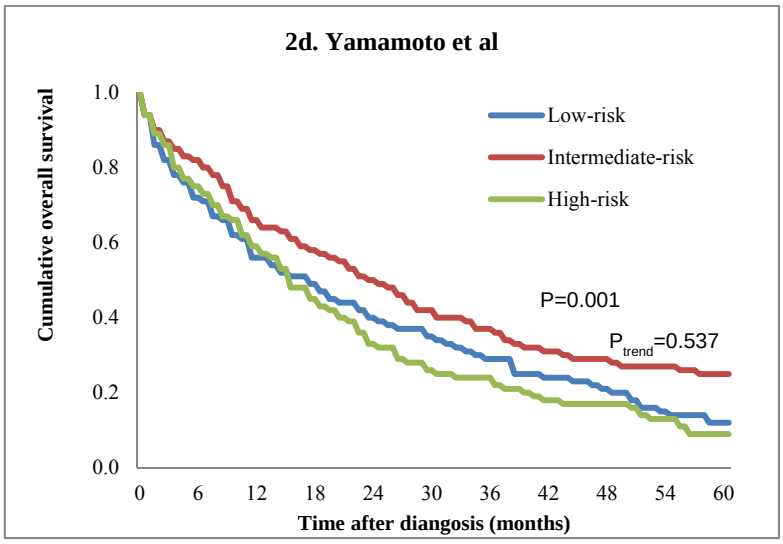
		(16.9%)	(7.1%)	(13.6%)	(13.4%)
	Liver only	22 (8.3%)	9 (5.8%)	20 (9.1%)	51 (7.9%)
	Brain only	5 (1.9%)	2 (1.3%)	2 (0.9%)	9 (1.4%)
	Soft tissue only	5 (1.9%)	0 (0.0%)	3 (1.4%)	8 (1.2%)
	Other organ only	2 (0.8%)	1 (0.6%)	3 (1.4%)	6 (0.9%)
	Multiple sites	118 (44.4%)	104 (66.7%)	106 (48.2%)	328 (51.1%)
	Unknown	12 (4.5%)	4 (2.6%)	10 (4.5%)	26 (4.0%)
Surgery	No surgery	155 (58.3%)	84 (53.8%)	165 (75.0%)	404 (62.9%)
	Mastectomy	111 (41.7%)	63 (40.4%)	51 (23.2%)	225 (35.0%)
	Breast conserving surgery	0 (0.0%)	9 (5.8%)	4 (1.8%)	13 (2.0%)
Chemotherapy	No	101 (38.0%)	77 (49.4%)	53 (24.1%)	231 (36.0%)
	Yes	164 (61.7%)	79 (50.6%)	94 (42.7%)	337 (52.5%)
	Unknown	1 (0.4%)	0 (0.0%)	73 (33.2%)	74 (11.5%)
Radiotherapy	No	115 (43.2%)	106 (67.9%)	129 (58.6%)	350 (54.5%)
	Yes	96 (36.1%)	45 (28.8%)	19 (8.6%)	160 (24.9%)
	Unknown	55 (20.7%)	5 (3.2%)	72 (32.7%)	132 (20.6%)
Hormone therapy	No	63 (23.7%)	95 (60.9%)	120 (54.5%)	278 (43.3%)
	Yes	121 (45.5%)	58 (37.2%)	29 (13.2%)	208 (32.4%)
	Unknown	82 (30.8%)	3 (1.9%)	71 (32.3%)	156 (24.3%)

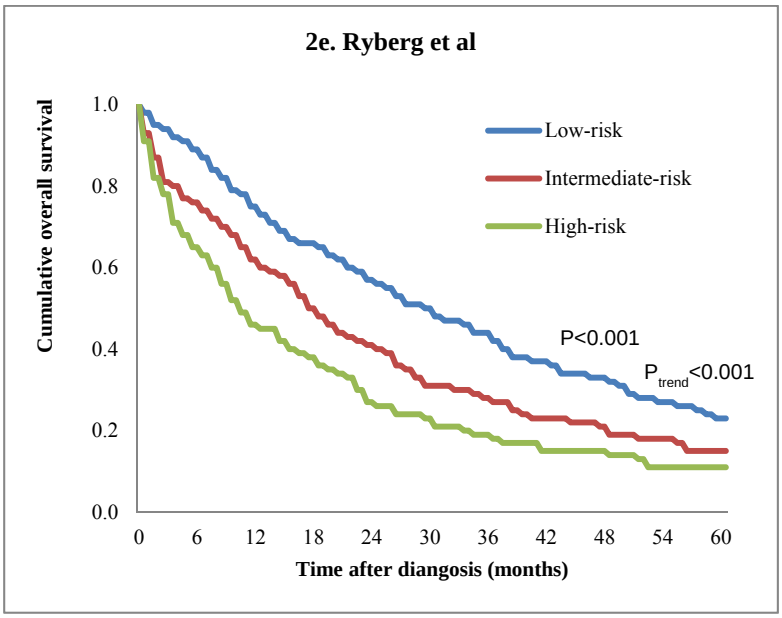
Figure 8.2 Kaplan-Meier survival curves of low, intermediate and high-risk groups. Risk groups were defined by tertiles of risk scores of prediction models for patients with *de novo* metastatic breast cancer.

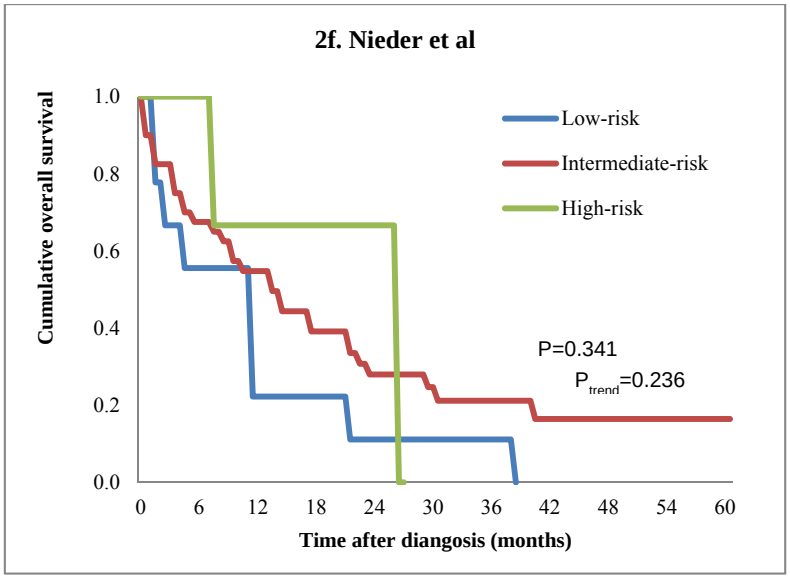


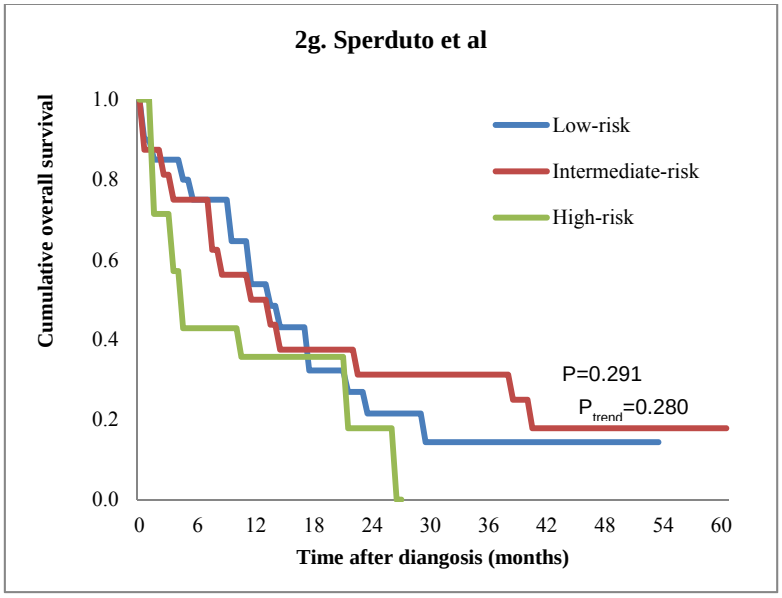


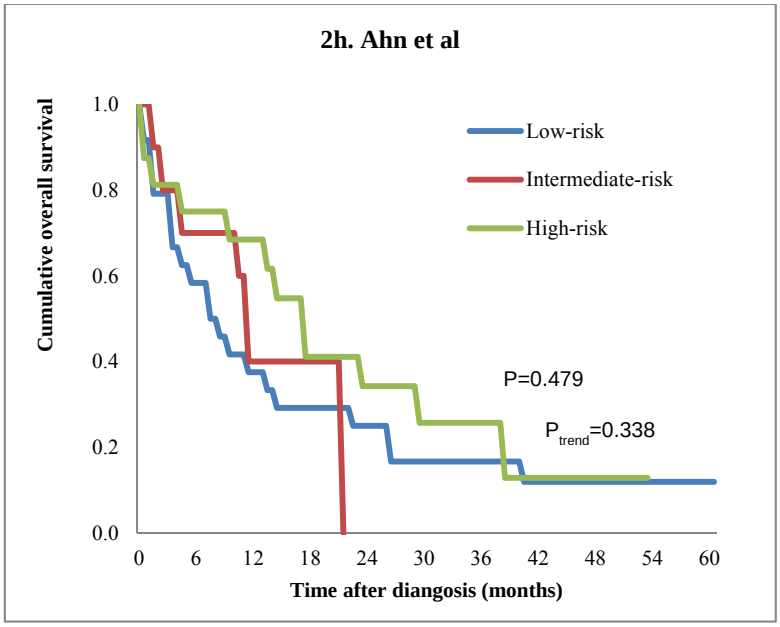

2c. Rabinovich et al
Cumulative overall survival
1.0
0.8
0.6
0.4
0.2
0.0
0
6
12
18
24
30
36
42
48
54
60
Time after diangosis (months)
Low-risk
High-risk
P<0.001
P trend <0.001


2d. Yamamoto et al
Cumulative overall survival
1.0
0.8
0.6
0.4
0.2
0.0
0
6
12
18
24
30
36
42
48
54
60
Time after diangosis (months)
Low-risk
Intermediate-risk
High-risk
P=0.001
P trend =0.537


2e. Ryberg et al
Cumulative overall survival
1.0
0.8
0.6
0.4
0.2
0.0
0
6
12
18
24
30
36
42
48
54
60
Time after diangosis (months)
Low-risk
Intermediate-risk
High-risk
P<0.001
P trend <0.001


2f. Nieder et al
Cumulative overall survival
1.0
0.8
0.6
0.4
0.2
0.0
0
6
12
18
24
30
36
42
48
54
60
Time after diangosis (months)
Low-risk
Intermediate-risk
High-risk
P=0.341
P trend=0.236


2g. Sperduto et al
Cumulative overall survival
1.0
0.8
0.6
0.4
0.2
0.0
0
6
12
18
24
30
36
42
48
54
60
Time after diangosis (months)
Low-risk
Intermediate-risk
High-risk
P=0.291
P trend=0.280


2h. Ahn et al
Cumulative overall survival
1.0
0.8
0.6
0.4
0.2
0.0
0
6
12
18
24
30
36
42
48
54
60
Time after diangosis (months)
Low-risk
Intermediate-risk
High-risk
P=0.479
P trend=0.338

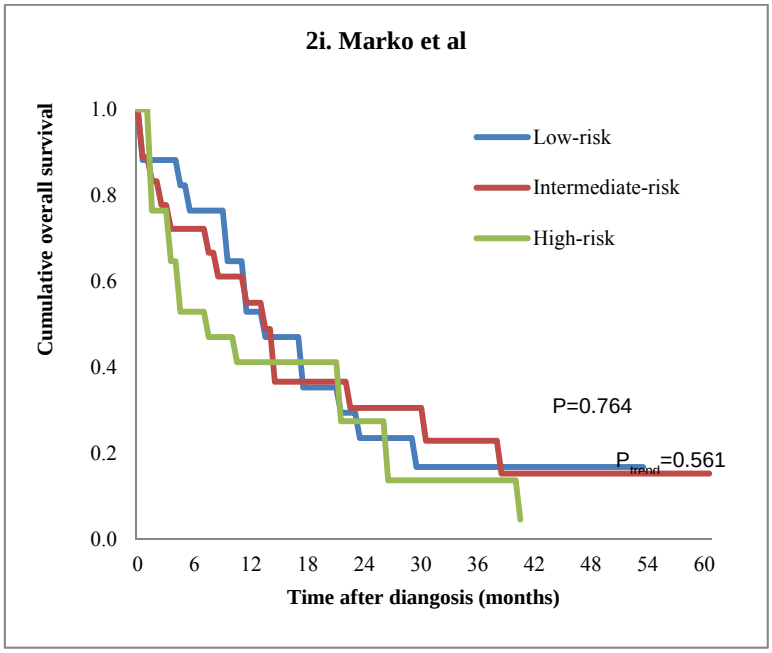
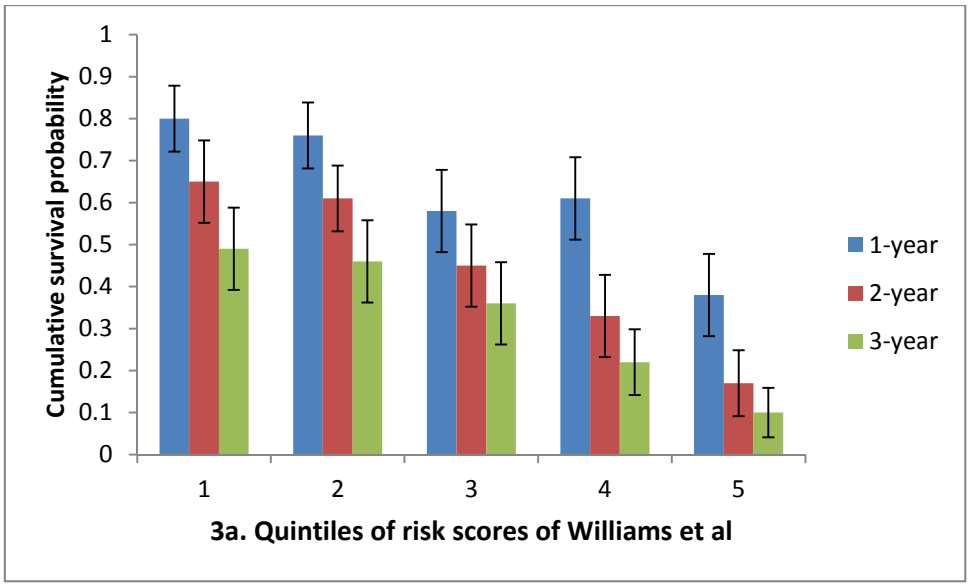

2i. Marko et al
Cumulative overall survival
1.0
0.8
0.6
0.4
0.2
0.0
0
6
12
18
24
30
36
42
48
54
60
Time after diangosis (months)
Low-risk
Intermediate-risk
High-risk
P=0.764
P trend =0.561

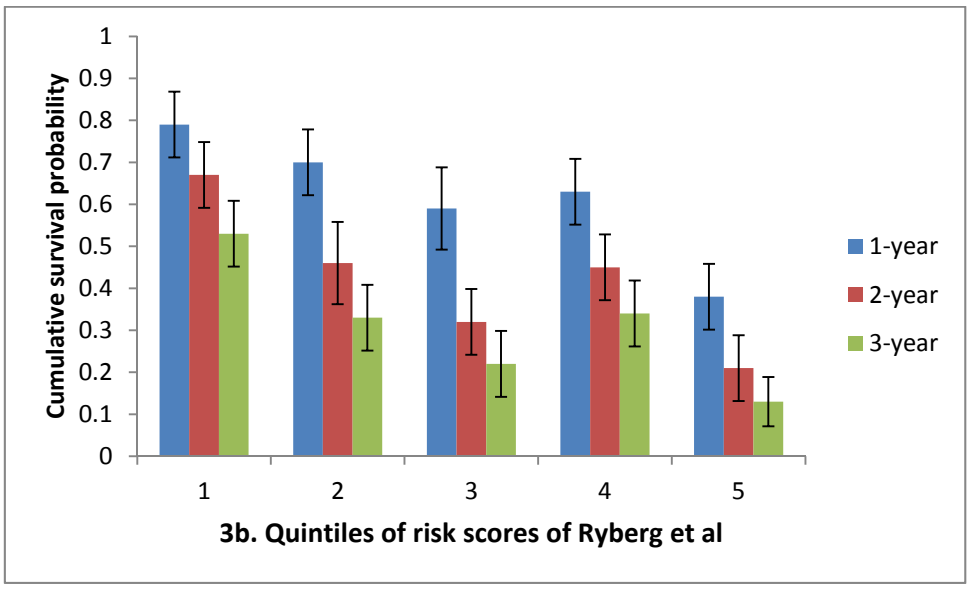
Table 8.4 Validation of selected models for prediction of survival of patients with *de novo* metastatic breast cancer

Model	Number of subjects available for validation	Possible range of scores	Observed range of scores	C-statistic (95% CI)
Nash et al.	642	0.23-3.44	0.23-3.44	0.51 (0.48,0.53)
Williams et al.	571[a]	-2.00-32.00	1.23-32.00	0.63 (0.60,0.66)[d]
Rabinovich et al.	642	0.80-2.38	0.80-1.05	0.55 (0.53,0.57)
Yamamoto et al.	642	0.00-6.33	3.33-6.33	0.50 (0.48,0.53)
Ryberg et al.	642	0.00-50.00	0.00-25.00	0.61 (0.59,0.64)
Nieder et al.	52[c]	0.00-5.00	1.00-3.00	0.55 (0.48,0.61)
Sperduto et al.	50[b,c]	0.00-4.00	1.50-4.00	0.56 (0.47,0.65)
Ahn et al.	50[b,c]	0.00-325.00	0.00-138.00	0.56 (0.46,0.66)
Marko et al.	52[c]	0.00-375.00	44.50-108.60	0.55 (0.45,0.64)

[a] Patients with brain metastases excluded
[b] Patients with equivocal HER2 status were excluded
[c] Exclusively patients with brain metastasis
[d] C-statistic for complete case analysis based on 297 patients was 0.63 (95% CI, 0.59-0.67)

Figure 8.3 1-, 2- and 3-year cumulative survival probability for different risk groups.Risk groups were defined by quintiles of risk scores of Williams et al.'s and Ryberg et al.'s model. 1st quintile is the group with the highest predicted survival probability and 5th quintile is with the lowest predicted survival probability.





8.4 Discussion

Survival after *de novo* metastatic breast cancer, a relatively common condition among breast cancer patients in South East Asia, varies considerably. In this study, we showed that this highly variable prognosis can be predicted using currently available prediction rules, only to a certain extent in Asian patients. Overall, the prediction performance in the present series in Asia was not as good as in the original reports. Some of these prediction rules, which were identified through systematic review of the literature, used easily available clinical information such as age, hormone receptor status and site of metastasis. Some other models included biomarkers, which are not routinely available during the work up of breast cancer patients such as CTC and LDH.

We validated nine of the models in our Asian dataset and found that two models performed moderately well. In fact, with basic clinical information, (i.e. grade, ER status and site of metastasis), these models were able to classify patients as high risk and low risk. Based on risk scores calculated from Williams et al.'s and Ryberg et al.'s models, which included simple freely available clinical information, the difference of 3-year survival probability between the highest and lowest quintiles was close to 40%. Still, there was substantial overlap between the categories, and the current prediction rules were at best fairly able to discriminate between low and high risk patients (highest C-statistic=0.63). Comparing to the other three models developed for all metastatic breast cancer patients, the models developed by Williams et al and Ryberg et al incorporated ER status and also grouped metastatic site into more categorizes. We were unable to validate the models which included advanced biomarkers, as this information was not routinely captured in our patients.

The inferior performance of the models in our Asian dataset as compared to the original report could be explained by unavailability of some predictors in our cohort and the fact that these indices/models were not specifically designed for *de novo* metastatic breast cancer. Another explanation could be that the Western derived models are not suitable for Asia setting. For example, in women with stage I-III breast cancer, Adjuvant!Online overpredicted survival by almost 7% and this overprediction was especially pronounced in younger women and women of Malay descent [197]. The underlying cause might be different distributions of age, tumor characteristics, competing risks and lifestyles factors. Several studies have reported that Asian breast cancer patients are more likely to be premenopausal, ER/PR-negative and HER2-positive [102, 365, 366]. Such differences could result in more skewed or more restricted range of prediction scores (Table 8.4).

Accuracy of predicting survival is crucial for women with *de novo* metastatic breast cancer as treatment varies widely, from no treatment at all, to removal of primary tumor and aggressive systemic treatment. The use of endocrine therapy and anti-HER2 drugs has been shown to prolong survival of metastatic patients [367-369]. Many randomized control trials have also reported significant survival benefit from modern chemotherapeutic agents, such as taxanes [370]. Recent studies have suggested that women who undergo surgery for *de novo* metastatic breast cancer have a significantly lower risk of death as compared to those who do not [145, 338, 339]. However the high proportion of patients not treated in our cohort or different response to treatment between Asian and Caucasian women may affect the usefulness of certain predictors such as hormone receptor status as well as the overall performance of the prediction models.

We acknowledge that our study suffers from limitations. The main limitation of the current study is the unavailability of certain clinical variables for prediction in our database such as performance status and LDH. Performance status, either recorded in Zubrod/ECOG/WHO or KPS, is a significant predictor in 11 indices/models. According to the development studies, 60% to 79% of their study population in fact had good performance status (Zubrod/ECOG/WHO= 0 or 1 or KPS ≥70). Based on the results from a subset of patients with comorbidity data in our validation set, our assumption of patients to be generally fit may have resulted in some overestimation of predicted survival probabilities for a subset of patients. The number of CTC has been shown to be highly predictive for overall survival in patients with metastatic breast cancer [371, 372]. The CELLSEARCH test (Veridex, LLC, Raritan, NJ, USA) is the first and only clinically validated, FDA-cleared system for CTC assessment [373, 374]. However it is not routinely measured in Asia and is unlikely to be measured in future in low and middle income countries. The underperformance of models developed for brain metastasis maybe partially caused by the exclusion of non-treated patients in the development study, the lack of largest brain metastasis dimension and trastuzumab use in our validation dataset. Another limitation of our validation is the incomplete data of certain predictors. The pattern of missingness suggested missing at random and thus imputation was a better and more reasonable option than complete case analysis. The C-statistic for Williams et al's model from complete case analysis of 297 patients with grade, ER status and metastatic site(s) was 0.63 (95% CI, 0.59-0.67), which was very similar to the result from imputation (0.63, 95% CI, 0.60-0.66). However the standard errors and confidence intervals of the estimates might be too low as we ignored the uncertainty of imputed values by single imputation.

We conclude that existing prognostic models can only moderately predict survival of women with *de novo* metastatic breast cancer in the Asian setting. New models derived from a representative sample from an Asian population with different disease burden, would be able to accurately discriminate between patients with relatively good versus poor prognosis better.

Chapter 9 Overall discussion and future perspectives

Clinicians and breast cancer patients have a keen interest in knowing the probability of survival and disease progression so that they can choose the most appropriate treatment, decide the frequency and intensity of post-treatment surveillance and plan for future activities. However there is no one-size-fits-all solution due to the heterogeneous nature of breast cancer. Although classification systems based on histopathological or molecular features have been established for risk stratification and tailored management plan, differences in response to treatment and disease courses have been observed among patients with similar tumor profiles. Without a clear picture of the biological and clinical diversity of breast cancer, we are nowhere near achieving individualized prognostication and treatment.

This thesis focuses on outcome and prognostic factors in understudied subgroups of breast cancer patients: male breast cancer patients, patients with breast carcinoma *in situ*, patients with de *novo* metastatic breast cancer, and patients from Southeast Asia. Comparison between male and female breast cancer, between diseased and healthy individuals in the first and second study helped us better understand breast cancer risk and progression. Study of periodic trend of breast cancer outcome provided evidence on effect of early detection and treatment advancement. Identification of clinically important prognostic factors such as young age and family history for women diagnosed with carcinoma *in situ* would improve evidence-based decision-making. Last but not least, validation of prognostic tools in Southeast Asia evaluated the clinical utility of these tools in this region and addressed pressing needs for region-specific tools.

9.1 Male breast cancer

Male breast cancer is a rare disease and treatment for male breast cancer typically follows the guidelines set for female breast cancer as it is considered resemble postmenopausal female breast cancer. The first study in this thesis found that the improvement in survival among male breast cancer over the last 30 years was not as pronounced as female patients. The poorer observed survival of male patients can be explained by older age at diagnosis, late stage and treatment differences. These findings suggest lack of early detection and specific treatment guidelines for men. Therefore it is important to raise public awareness of this rare disease and educate men on the early symptoms of breast cancer. Studies have reported higher proportion of hormone receptor positive among male patients and mixed results on HER2 status. However men with hormone receptor positive disease have reportedly poorer adherence to tamoxifen with 20% patients discontinuing due to side effects [375]. Most of these studies suffered from small sample size and variation in study designs and methods [376]. Since men are diagnosed with older median age, age, comorbidity and performance status should be carefully considered when determine chemotherapy with toxic adverse effects. In addition, data on treatment of male breast cancer is mainly from small retrospective and single institutional studies. Conducting RCT for male breast cancer is very complicated due to low incidence. Unfortunately detailed treatment data such as compliance and regimes are not available in population-based cancer registries. Multi-institutional collaboration is crucial to understand tumor biology, effect of treatment and compliance to treatment, especially hormonal therapy on male breast cancer. The ultimate goal is to formulate guidelines and protocols for treatment and

surveillance based on robust and consistent observations from large scaled prospective studies.

9.2 Breast carcinoma *in situ*

The risk of subsequent invasive breast cancer varies between patients as breast carcinoma *in situ* can either be an indolent non-progressive lesion which only requires close monitoring, or a precursor to invasive breast cancer which should be treated more aggressively using mastectomy or chemotherapy [377]. However carcinoma *in situ* patients are more likely to overestimate their risk of recurrence than early invasive breast cancer parents due to lack awareness about their prognosis [378]. This wrong self-perception on recurrence will eventually lead to anxiety and influence decision on treatment and long-term follow-up plan. In the second study, we found that, among patients diagnosed with breast carcinoma *in situ,* family history increased the risk for a contralateral invasive breast cancer. Also women who diagnosed with breast carcinoma *in situ* before the age of 40 years were at higher risk of subsequent invasive breast cancer. Since mastectomy could lower the risk of local recurrence than breast conserving surgery, option of mastectomy should be discussed with young patients with family history. Meanwhile for patients with low risk of invasive event, overtreatment should be avoided. There is lack of evidence on best post-treatment surveillance procedure for *in situ* cases. A few recommendations have been proposed ranging from physical examination every 6 to 12 months to a mammogram every 6 to 12 months, especially during the first year of diagnosis and may vary by type of surgery received [379]. However these recommendations did not take young age at diagnosis and family history into consideration. Findings from our study indicate

that patients with these risk factors should be monitored closely and continuously as the risk persists even 15 years after diagnosis. For patients with low risk breast carcinoma *in situ*, there is an increasing interest in a watchful waiting or active surveillance approach which has been offered to prostate cancer patients to avoid treatment [380, 381]. There are two ongoing RCTs (LORD trial and LORIS trial) in the Netherlands and United Kingdom to compare the effect of active surveillance with the standard treatment for breast carcinoma *in situ* patients.

The second study was conducted in Swedish population and its implication in Asia is uncertain. Singapore is the first country to establish nationwide screening programme (Breast Screen Singapore) in Asia since January 2002 [382]. In Singapore breast screening pilot project, 26% of all screen-detected cancers were carcinoma *in situ*, which is much higher than 10–15% of cancers detected in women not invited for screening [383, 384]. With rising breast cancer incidence in Singapore and more women undergoing mammographic screening, *in situ* cancer is poised to become more common. It is important to understand how introduction of screening could affect tumor characteristics and outcome of *in situ* disease in Singapore. Similarities and differences in the disease burden, pattern of presentation and outcome between ethnic groups in Asia and between Asian and Caucasian populations have not yet been studied.

9.3 Prognostication for Southeast Asian patients

In the past decades, we have witnessed a growing interest in, and use of, prognostic models in clinical decision-making process, especially in western countries. However validation studies of Adjuvant! Online conducted in Asia and validation of CancerMath conducted in this thesis have showed overestimation of prognosis

in this region by these prediction tools. Existing models for metastatic breast cancer performed unsatisfactorily in this region as presented in the fourth study in this thesis. Many of these models were developed more than a decade ago and the more recent models included advanced biomarkers such as circulating tumor cell which is not obtained during routine clinical practice in Asia. Studies have revealed that breast cancer in Asian women is distinctive from their western counterparts in many perspectives such as underlying risk factors, disease presentation, tumor characteristics, lifestyle after cancer, treatment response and tolerance to side effects. The breast cancer epidemic in Asia calls for region-specific tools to improve outcome and treatment prediction.

In order to perform excellent research in the field of prognostication, complete, accurate and consistent longitudinal data collection is crucial, but this is not very easy to find (yet) in Asia. The incidence rate of breast cancer in Singapore is among the highest in Asia and is steeply rising. The availability of national and institutional cancer registries in combination with the multi-ethnic build-up of the society makes Singapore unique for validation of established prediction models and development of new models in an Asian setting. Multi-institutional effort to set up the Singapore Malaysia Hospital Based Breast Cancer Registry was one of the initiatives to facilitate clinical research on breast cancer in Southeast Asia. Thus far published and ongoing studies from this collaboration include validation of prognostic factors and models, comparison of outcome between ethnic groups and countries, evaluation of treatment effect and toxicity. We anticipate more hospitals in this region to participate and more comprehensive data to be available for future studies.

Missing data and incomplete follow up are common problems in many retrospective clinical databases and become major obstacles for validation and development of prognostic models. A substantial number of cases with follow-up less than 5 years were excluded in most analyses in study 3. In study 4, we applied missing data imputation based on the assumption of missing at random. Both approaches have their limitations, which may reduce the power of the study and potentially cause bias (reliability and validity of the findings). This issue addresses the importance of prospective collection of data primarily for research purpose with clear definition and accurate measurement of outcome and covariates.

9.4 Public health implications

Prognostic and survivorship research for breast cancer is not highly recognized for its impact on public health as it focuses on tertiary prevention and only targets individuals affected by breast cancer. Currently there are nearly 2.8 million women with a history of breast cancer living in the United States [385]. High prevalence of cancer survivors is associated with notable economic burden to the society, directly from increased medical expenditures and indirectly from productivity loss [386]. Given the increasing trend of breast cancer incidence and improvement in survival, Asian countries will expect more women living with the disease in the population. Besides financial impact, risk of recurrence and secondary cancer, reduction in quality of life and wellbeing of family members and caregivers will soon become major public health concerns.

Current public health efforts on cancer control such as promoting healthy lifestyles can be adapted for cancer survivors for early detection and prevention of recurrence and secondary cancer. Prognostic indicator or models can help us identify targeted

groups with high risk and formulate more personalized approach to deliver medical advice and care.

Prolonging life is considered as a main goal when we evaluate effect of treatment or other interventions. For high-income countries like Singapore with better healthcare facilities and cancer treatment but aging population, patient-reported outcome, which includes symptoms, functioning, health related quality of life and satisfaction with care after diagnosis and treatment becomes an important outcome measure in addition to survival. International efforts have been made in recent decades to develop patient-reported outcome measures (PROMs) for various types of cancer and establish guidance related to PROMs [387-390]. However robust and standardized PROM instruments specifically designed for breast cancer patients in Singapore is not yet available. Similar to prognostic models, many assessment tools were designed using Western population and the parameters included may not correctly address the health concerns of Asian patients. Future studies on PROMs for breast cancer patients in Asia are needed.

In contrast to high income countries with government-funded screening programme and medical subsidy for treatment, delayed presentation and poor compliance with treatment remain major concerns in many developing countries in Asia. As a result, significant survival disparities between countries and between socioeconomic groups have been highlighted in many studies. Lack of awareness of breast cancer, poor access to quality healthcare facilities and health-seeking behaviour associated with cultural beliefs are the main barriers to better outlook for women with breast cancer. The high cost and some technical limitation have hindered adaption of certain biomarker testing such as fluorescence in situ hybridization test for HER2, microarray based gene expression, circulating tumor cell, etc in local clinical setting.

Willingness and openness to discussion regarding life expectancy and survival probability in fact varies from patient to patient. We need to consider these facts when developing new region-specific prognostic tool.

Moving forward, the utility and cost effectiveness of prognostic markers or models should be evaluated using local data before adaption or making any modification. Ongoing data collection and biospecimen banking such as effort made by the Singapore Breast Cancer Cohort Study, which recruits existing and newly diagnosed breast cancer patients at five public hospitals in Singapore, should be extended to other parts of the region to support large scaled clinical and translational research to improve quality of care.

Bibliography

1. *Latest world cancer statistics. Global cancer burden rises to 14.1 million new cases in 2012: Marked increase in breast cancers must be addressed*. 2013, International Agency for Research on Cancer (IARC), WHO. p. 1.
2. Ferlay, J., et al., *Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008.* Int J Cancer, 2010. **127**(12): p. 2893-917.
3. *Trend in cancer incidence in singapore 1968-2007*. 2010, Singapore Cancer Registry.
4. Ferlay J, S.I., Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F., *GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]*. 2013: Lyon, France.
5. Yip, C.H., *Breast Cancer in Asia*, in *Methods in Molecular Biology, Cancer Epidemiology* M. Verma, Editor., humana Press.
6. Bhoo-Pathy, N., et al., *Breast cancer research in Asia: adopt or adapt Western knowledge?* Eur J Cancer, 2013. **49**(3): p. 703-9.
7. Berry, D.A., et al., *Effect of screening and adjuvant therapy on mortality from breast cancer.* N Engl J Med, 2005. **353**(17): p. 1784-92.
8. Chia, K.S., et al., *Profound changes in breast cancer incidence may reflect changes into a Westernized lifestyle: a comparative population-based study in Singapore and Sweden.* Int J Cancer, 2005. **113**(2): p. 302-6.
9. *Trends of female breast cancer in Singapore 2006-2010*, in *Health Factsheet*. 2012, National Registry of Disease Office: Singapore.
10. La Vecchia, C., F. Levi, and F. Lucchini, *Descriptive epidemiology of male breast cancer in Europe.* Int J Cancer, 1992. **51**(1): p. 62-6.
11. Giordano, S.H., et al., *Breast carcinoma in men: a population-based study.* Cancer, 2004. **101**(1): p. 51-7.
12. Kreiter, E., et al., *Breast cancer: trends in international incidence in men and women.* Br J Cancer, 2014. **110**(7): p. 1891-7.
13. Clemmesen, J., *Carcinoma of the breast; results from statistical research.* Br J Rad, 1948. **21**: p. 583-90.
14. Verkasalo, P.K., et al., *Circulating levels of sex hormones and their relation to risk factors for breast cancer: a cross-sectional study in 1092 pre- and postmenopausal women (United Kingdom).* Cancer Causes Control, 2001. **12**(1): p. 47-59.
15. Stavraky, K. and S. Emmons, *Breast cancer in premenopausal and postmenopausal women.* J Natl Cancer Inst, 1974. **53**(3): p. 647-54.
16. Jatoi, I., W.F. Anderson, and P.S. Rosenberg, *Qualitative age-interactions in breast cancer: a tale of two diseases?* Am J Clin Oncol, 2008. **31**(5): p. 504-6.
17. de Waard, F., *Premenopausal and postmenopausal breast cancer: one disease or two?* J Natl Cancer Inst, 1979. **63**(3): p. 549-52.
18. Lipnick, R., et al., *A case-control study of risk indicators among women with premenopausal and early postmenopausal breast cancer.* Cancer, 1984. **53**(4): p. 1020-4.
19. Hirose, K., et al., *Association of family history and other risk factors with breast cancer risk among Japanese premenopausal and postmenopausal women.* Cancer Causes Control, 2001. **12**(4): p. 349-58.

20. Walker, K., D.J. Bratton, and C. Frost, *Premenopausal endogenous oestrogen levels and breast cancer risk: a meta-analysis.* Br J Cancer, 2011. **105**(9): p. 1451-7.
21. Huang, Z., et al., *Dual effects of weight and weight gain on breast cancer risk.* JAMA, 1997. **278**(17): p. 1407-11.
22. Gabriel, C.A. and S.M. Domchek, *Breast cancer in young women.* Breast Cancer Res, 2010. **12**(5): p. 212.
23. Anders, C.K., et al., *Young age at diagnosis correlates with worse prognosis and defines a subset of breast cancers with shared patterns of gene expression.* J Clin Oncol, 2008. **26**(20): p. 3324-30.
24. Leong, S.P., et al., *Is breast cancer the same disease in Asian and Western countries?* World J Surg, 2010. **34**(10): p. 2308-24.
25. Anderson, W.F., et al., *Shifting breast cancer trends in the United States.* J Clin Oncol, 2007. **25**(25): p. 3923-9.
26. Hausauer, A.K., et al., *Recent breast cancer trends among Asian/Pacific Islander, Hispanic, and African-American women in the US: changes by tumor subtype.* Breast Cancer Res, 2007. **9**(6): p. R90.
27. Phipps, A.I. and C.I. Li, *Breast Cancer Biology and Clinical Characteristics*, in *Breast Cancer Epidemiology*, C.I. Li, Editor. 2010, Springer Science+Business Media. p. 21.
28. Pharoah, P.D., et al., *Family history and the risk of breast cancer: a systematic review and meta-analysis.* Int J Cancer, 1997. **71**(5): p. 800-9.
29. Lichtenstein, P., et al., *Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland.* N Engl J Med, 2000. **343**(2): p. 78-85.
30. Ghoussaini, M., P.D. Pharoah, and D.F. Easton, *Inherited genetic susceptibility to breast cancer: the beginning of the end or the end of the beginning?* Am J Pathol, 2013. **183**(4): p. 1038-51.
31. Antoniou, A.C., et al., *Breast-cancer risk in families with mutations in PALB2.* N Engl J Med, 2014. **371**(6): p. 497-506.
32. Maxwell, K.N. and K.L. Nathanson, *Common breast cancer risk variants in the post-COGS era: a comprehensive review.* Breast Cancer Res, 2013. **15**(6): p. 212.
33. Michailidou, K., et al., *Large-scale genotyping identifies 41 new loci associated with breast cancer risk.* Nat Genet, 2013. **45**(4): p. 353-61, 361e1-2.
34. Boyd, N.F., et al., *Mammographic density and the risk and detection of breast cancer.* N Engl J Med, 2007. **356**(3): p. 227-36.
35. Pettersson, A., et al., *Mammographic Density Phenotypes and Risk of Breast Cancer: A Meta-analysis.* J Natl Cancer Inst, 2014. **106**(5).
36. van Gils, C.H., et al., *Changes in mammographic breast density and concomitant changes in breast cancer risk.* Eur J Cancer Prev, 1999. **8**(6): p. 509-15.
37. Kerlikowske, K., et al., *Longitudinal measurement of clinical mammographic breast density to improve estimation of breast cancer risk.* J Natl Cancer Inst, 2007. **99**(5): p. 386-95.
38. Cancer, C.G.o.H.F.i.B., *Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies.* The Lancet Oncology, 2012. **13**(11): p. 1141-1151.

39. Kelsey, J.L., M.D. Gammon, and E.M. John, *Reproductive factors and breast cancer.* Epidemiol Rev, 1993. **15**(1): p. 36-47.
40. Lipworth, L., L.R. Bailey, and D. Trichopoulos, *History of breast-feeding in relation to breast cancer risk: a review of the epidemiologic literature.* J Natl Cancer Inst, 2000. **92**(4): p. 302-12.
41. Rossouw, J.E., et al., *Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial.* JAMA, 2002. **288**(3): p. 321-33.
42. Beral, V. and C. Million Women Study, *Breast cancer and hormone-replacement therapy in the Million Women Study.* Lancet, 2003. **362**(9382): p. 419-27.
43. Verkooijen, H.M., et al., *The incidence of breast cancer and changes in the use of hormone replacement therapy: a review of the evidence.* Maturitas, 2009. **64**(2): p. 80-5.
44. Ravdin, P.M., et al., *The decrease in breast-cancer incidence in 2003 in the United States.* N Engl J Med, 2007. **356**(16): p. 1670-4.
45. Barton, M.B., R. Harris, and S.W. Fletcher, *The rational clinical examination. Does this patient have breast cancer? The screening clinical breast examination: should it be done? How?* JAMA, 1999. **282**(13): p. 1270-80.
46. Bancej, C., et al., *Original Paper: Contribution of clinical breast examination to mammography screening in the early detection of breast cancer.* Journal of Medical Screening, 2003. **10**(1): p. 16-21.
47. Kosters, J.P. and P.C. Gotzsche, *Regular self-examination or clinical examination for early detection of breast cancer.* Cochrane Database Syst Rev, 2003(2): p. CD003373.
48. Dowling, E.C., et al., *Breast and cervical cancer screening programme implementation in 16 countries.* J Med Screen, 2010. **17**(3): p. 139-46.
49. Woo, P.P., J.J. Kim, and G.M. Leung, *What is the most cost-effective population-based cancer screening program for Chinese women?* J Clin Oncol, 2007. **25**(6): p. 617-24.
50. Okonkwo, Q.L., et al., *Breast cancer screening policies in developing countries: a cost-effectiveness analysis for India.* J Natl Cancer Inst, 2008. **100**(18): p. 1290-300.
51. Kang, M.H., et al., *The National Cancer Screening Program for breast cancer in the Republic of Korea: is it cost-effective?* Asian Pac J Cancer Prev, 2013. **14**(3): p. 2059-65.
52. Lui, C.Y., et al., *Opportunistic breast cancer screening in Hong Kong; a revisit of the Kwong Wah Hospital experience.* Hong Kong Med J, 2007. **13**(2): p. 106-13.
53. Dahlui, M., S. Ramli, and A.M. Bulgiba, *Breast cancer prevention and control programs in Malaysia.* Asian Pac J Cancer Prev, 2011. **12**(6): p. 1631-4.
54. Bihrmann, K., et al., *Performance of systematic and non-systematic ('opportunistic') screening mammography: a comparative study from Denmark.* J Med Screen, 2008. **15**(1): p. 23-6.
55. de Gelder, R., et al., *Cost-effectiveness of opportunistic versus organised mammography screening in Switzerland.* Eur J Cancer, 2009. **45**(1): p. 127-38.

56. Ouedraogo, S., et al., *Breast cancer screening programmes: Challenging the coexistence with opportunistic mammography.* Patient Educ Couns, 2014. **97**(3): p. 410-7.
57. Independent, U.K.P.o.B.C.S., *The benefits and harms of breast cancer screening: an independent review.* Lancet, 2012. **380**(9855): p. 1778-86.
58. Gotzsche, P.C. and M. Nielsen, *Screening for breast cancer with mammography.* Cochrane Database Syst Rev, 2011(1): p. CD001877.
59. Esserman, L., Y. Shieh, and I. Thompson, *Rethinking screening for breast cancer and prostate cancer.* JAMA, 2009. **302**(15): p. 1685-92.
60. Miller, A.B., *Implications of the canadian national breast screening study.* Womens Health (Lond Engl), 2014. **10**(4): p. 345-7.
61. Miller, A.B., et al., *Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial.* BMJ, 2014. **348**: p. g366.
62. Pace, L.E. and N.L. Keating, *A systematic assessment of benefits and risks to guide breast cancer screening decisions.* JAMA, 2014. **311**(13): p. 1327-35.
63. Brewer, N.T., T. Salz, and S.E. Lillie, *Systematic review: the long-term effects of false-positive mammograms.* Ann Intern Med, 2007. **146**(7): p. 502-10.
64. Alamo-Junquera, D., et al., *Effect of false-positive results on reattendance at breast cancer screening programmes in Spain.* Eur J Public Health, 2012. **22**(3): p. 404-8.
65. Setz-Pels, W., et al., *Re-attendance after false-positive screening mammography: a population-based study in the Netherlands.* Br J Cancer, 2013. **109**(8): p. 2044-50.
66. Centre, N.B.C., *Breast fine needle aspiration cytology and core biopsy: a guide for practice*. First ed. 2004, Camperdown, NSW, Australia: National Breast Cancer Centre.
67. Le Gal, M., et al., *Mammographic features of 455 invasive lobular carcinomas.* Radiology, 1992. **185**(3): p. 705-8.
68. De Leeuw, W.J., et al., *Simultaneous loss of E-cadherin and catenins in invasive lobular breast cancer and lobular carcinoma in situ.* J Pathol, 1997. **183**(4): p. 404-11.
69. Wasif, N., et al., *Invasive lobular vs. ductal breast cancer: a stage-matched comparison of outcomes.* Ann Surg Oncol, 2010. **17**(7): p. 1862-9.
70. Pestalozzi, B.C., et al., *Distinct clinical and prognostic features of infiltrating lobular carcinoma of the breast: combined results of 15 International Breast Cancer Study Group clinical trials.* J Clin Oncol, 2008. **26**(18): p. 3006-14.
71. Lamovec, J. and M. Bracko, *Metastatic pattern of infiltrating lobular carcinoma of the breast: an autopsy study.* J Surg Oncol, 1991. **48**(1): p. 28-33.
72. Arpino, G., et al., *Infiltrating lobular carcinoma of the breast: tumor characteristics and clinical outcome.* Breast Cancer Res, 2004. **6**(3): p. R149-56.
73. Yerushalmi, R., M.M. Hayes, and K.A. Gelmon, *Breast carcinoma--rare types: review of the literature.* Ann Oncol, 2009. **20**(11): p. 1763-70.

74. Li, C.I., D.J. Uribe, and J.R. Daling, *Clinical characteristics of different histologic types of breast cancer.* Br J Cancer, 2005. **93**(9): p. 1046-52.
75. Bloom, H.J. and W.W. Richardson, *Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years.* Br J Cancer, 1957. **11**(3): p. 359-77.
76. Elston, C.W. and I.O. Ellis, *Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up.* Histopathology, 1991. **19**(5): p. 403-10.
77. Rakha, E.A., et al., *Prognostic significance of Nottingham histologic grade in invasive breast carcinoma.* J Clin Oncol, 2008. **26**(19): p. 3153-8.
78. Grotenhuis, B.A., T.M. Klem, and W.W. Vrijland, *Treatment outcome in breast cancer patients with ipsilateral supraclavicular lymph node metastasis at time of diagnosis: a review of the literature.* Eur J Surg Oncol, 2013. **39**(3): p. 207-12.
79. Brito, R.A., et al., *Long-term results of combined-modality therapy for locally advanced breast cancer with ipsilateral supraclavicular metastases: The University of Texas M.D. Anderson Cancer Center experience.* J Clin Oncol, 2001. **19**(3): p. 628-33.
80. Olivotto, I.A., et al., *Long-term survival of patients with supraclavicular metastases at diagnosis of breast cancer.* J Clin Oncol, 2003. **21**(5): p. 851-4.
81. Soerjomataram, I., et al., *An overview of prognostic factors for long-term survivors of breast cancer.* Breast Cancer Res Treat, 2008. **107**(3): p. 309-30.
82. Narod, S.A., *Tumour size predicts long-term survival among women with lymph node-positive breast cancer.* Curr Oncol, 2012. **19**(5): p. 249-53.
83. Johnson, R.H., F.L. Chien, and A. Bleyer, *Incidence of breast cancer with distant involvement among women in the United States, 1976 to 2009.* JAMA, 2013. **309**(8): p. 800-5.
84. Farley, T.A. and J.T. Flannery, *Late-stage diagnosis of breast cancer in women of lower socioeconomic status: public health implications.* Am J Public Health, 1989. **79**(11): p. 1508-12.
85. *Tamoxifen for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group.* Lancet, 1998. **351**(9114): p. 1451-67.
86. Fisher, B., et al., *Relative worth of estrogen or progesterone receptor and pathologic characteristics of differentiation as indicators of prognosis in node negative breast cancer patients: findings from National Surgical Adjuvant Breast and Bowel Project Protocol B-06.* J Clin Oncol, 1988. **6**(7): p. 1076-87.
87. Bardou, V.J., et al., *Progesterone receptor status significantly improves outcome prediction over estrogen receptor status alone for adjuvant endocrine therapy in two large breast cancer databases.* J Clin Oncol, 2003. **21**(10): p. 1973-9.
88. Dunnwald, L.K., M.A. Rossing, and C.I. Li, *Hormone receptor status, tumor characteristics, and prognosis: a prospective cohort of breast cancer patients.* Breast Cancer Res, 2007. **9**(1): p. R6.
89. Hammond, M.E., et al., *American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for*

immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Clin Oncol, 2010. **28**(16): p. 2784-95.
90. Winstanley, J., et al., *The long term prognostic significance of c-erbB-2 in primary breast cancer.* Br J Cancer, 1991. **63**(3): p. 447-50.
91. Borg, A., et al., *HER-2/neu amplification predicts poor survival in node-positive breast cancer.* Cancer Res, 1990. **50**(14): p. 4332-7.
92. Seol, H., et al., *Intratumoral heterogeneity of HER2 gene amplification in breast cancer: its clinicopathological significance.* Mod Pathol, 2012. **25**(7): p. 938-48.
93. Slamon, D.J., et al., *Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene.* Science, 1987. **235**(4785): p. 177-82.
94. Hanna, W.M., et al., *HER2 in situ hybridization in breast cancer: clinical implications of polysomy 17 and genetic heterogeneity.* Mod Pathol, 2014. **27**(1): p. 4-18.
95. Moja, L., et al., *Trastuzumab containing regimens for early breast cancer.* Cochrane Database Syst Rev, 2012. **4**: p. CD006243.
96. Hudis, C.A., *Trastuzumab--mechanism of action and use in clinical practice.* N Engl J Med, 2007. **357**(1): p. 39-51.
97. Wolff, A.C., et al., *Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update.* J Clin Oncol, 2013. **31**(31): p. 3997-4013.
98. Parise, C.A. and V. Caggiano, *Breast Cancer Survival Defined by the ER/PR/HER2 Subtypes and a Surrogate Classification according to Tumor Grade and Immunohistochemical Biomarkers.* J Cancer Epidemiol, 2014. **2014**: p. 469251.
99. Voduc, K.D., et al., *Breast cancer subtypes and the risk of local and regional relapse.* J Clin Oncol, 2010. **28**(10): p. 1684-91.
100. Brouckaert, O., et al., *Update on triple-negative breast cancer: prognosis and management strategies.* Int J Womens Health, 2012. **4**: p. 511-20.
101. Bauer, K.R., et al., *Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry.* Cancer, 2007. **109**(9): p. 1721-8.
102. Telli, M.L., et al., *Asian ethnicity and breast cancer subtypes: a study from the California Cancer Registry.* Breast Cancer Res Treat, 2011. **127**(2): p. 471-8.
103. Jack, R.H., et al., *Differences in breast cancer hormone receptor status in ethnic groups: a London population.* Eur J Cancer, 2013. **49**(3): p. 696-702.
104. Schnitt, S.J., *Classification and prognosis of invasive breast cancer: from morphology to molecular taxonomy.* Mod Pathol, 2010. **23 Suppl 2**: p. S60-4.
105. van 't Veer, L.J., et al., *Gene expression profiling predicts clinical outcome of breast cancer.* Nature, 2002. **415**(6871): p. 530-6.
106. Paik, S., et al., *A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer.* N Engl J Med, 2004. **351**(27): p. 2817-26.

107. Kroman, N., et al., *Factors influencing the effect of age on prognosis in breast cancer: population based study.* BMJ, 2000. **320**(7233): p. 474-8.
108. Nixon, A.J., et al., *Relationship of patient age to pathologic features of the tumor and prognosis for patients with stage I or II breast cancer.* J Clin Oncol, 1994. **12**(5): p. 888-94.
109. Fowble, B.L., et al., *The influence of young age on outcome in early stage breast cancer.* Int J Radiat Oncol Biol Phys, 1994. **30**(1): p. 23-33.
110. Han, W., et al., *Young age: an independent risk factor for disease-free survival in women with operable breast cancer.* BMC Cancer, 2004. **4**: p. 82.
111. Assi, H.A., et al., *Epidemiology and prognosis of breast cancer in young women.* J Thorac Dis, 2013. **5**(Suppl 1): p. S2-8.
112. Han, W., S.Y. Kang, and S. Korean Breast Cancer, *Relationship between age at diagnosis and outcome of premenopausal breast cancer: age less than 35 years is a reasonable cut-off for defining young age-onset breast cancer.* Breast Cancer Res Treat, 2010. **119**(1): p. 193-200.
113. de la Rochefordiere, A., et al., *Age as prognostic factor in premenopausal breast carcinoma.* Lancet, 1993. **341**(8852): p. 1039-43.
114. Yancik, R., et al., *Effect of age and comorbidity in postmenopausal breast cancer patients aged 55 years and older.* JAMA, 2001. **285**(7): p. 885-92.
115. van de Water, W., et al., *Association between age at diagnosis and disease-specific mortality among postmenopausal women with hormone receptor-positive breast cancer.* JAMA, 2012. **307**(6): p. 590-7.
116. Chia, K.S., et al., *Do younger female breast cancer patients have a poorer prognosis? Results from a population-based survival analysis.* Int J Cancer, 2004. **108**(5): p. 761-5.
117. Adami, H.O., et al., *The relation between survival and age at diagnosis in breast cancer.* N Engl J Med, 1986. **315**(9): p. 559-63.
118. Fredholm, H., et al., *Breast cancer in young women: poor survival despite intensive treatment.* PLoS One, 2009. **4**(11): p. e7695.
119. Patnaik, J.L., et al., *The influence of comorbidities on overall survival among older women diagnosed with breast cancer.* J Natl Cancer Inst, 2011. **103**(14): p. 1101-11.
120. Louwman, W.J., et al., *Less extensive treatment and inferior prognosis for breast cancer patient with comorbidity: a population-based study.* Eur J Cancer, 2005. **41**(5): p. 779-85.
121. Land, L.H., et al., *Comorbidity and survival after early breast cancer. A review.* Crit Rev Oncol Hematol, 2012. **81**(2): p. 196-205.
122. Braithwaite, D., et al., *Prognostic impact of comorbidity among long-term breast cancer survivors: results from the LACE study.* Cancer Epidemiol Biomarkers Prev, 2012. **21**(7): p. 1115-25.
123. Patterson, R.E., et al., *Medical comorbidities predict mortality in women with a history of early stage breast cancer.* Breast Cancer Res Treat, 2010. **122**(3): p. 859-65.
124. Nechuta, S., et al., *Comorbidities and breast cancer survival: a report from the Shanghai Breast Cancer Survival Study.* Breast Cancer Res Treat, 2013. **139**(1): p. 227-35.
125. Patnaik, J.L., et al., *Cardiovascular disease competes with breast cancer as the leading cause of death for older females diagnosed with breast*

cancer: a retrospective cohort study. Breast Cancer Res, 2011. **13**(3): p. R64.
126. Scott-Conner, C.E., et al., *An analysis of male and female breast cancer treatment and survival among demographically identical pairs of patients.* Surgery, 1999. **126**(4): p. 775-80; discussion 780-1.
127. Anderson, W.F., et al., *Male breast cancer: a population-based comparison with female breast cancer.* J Clin Oncol, 2010. **28**(2): p. 232-9.
128. Sabel, M.S., *Locoregional therapy of breast cancer: maximizing control, minimizing morbidity.* Expert Rev Anticancer Ther, 2006. **6**(9): p. 1281-99.
129. Clarke, M., et al., *Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: an overview of the randomised trials.* Lancet, 2005. **366**(9503): p. 2087-106.
130. Fisher, B., et al., *Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer.* N Engl J Med, 2002. **347**(16): p. 1233-41.
131. Veronesi, U., et al., *Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer.* N Engl J Med, 2002. **347**(16): p. 1227-32.
132. Poggi, M.M., et al., *Eighteen-year results in the treatment of early breast carcinoma with mastectomy versus breast conservation therapy: the National Cancer Institute Randomized Trial.* Cancer, 2003. **98**(4): p. 697-702.
133. Early Breast Cancer Trialists' Collaborative, G., et al., *Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials.* Lancet, 2011. **378**(9804): p. 1707-16.
134. Benson, J.R., et al., *Early breast cancer.* Lancet, 2009. **373**(9673): p. 1463-79.
135. Ebctcg Early Breast Cancer Trialists' Collaborative, G., *Effect of radiotherapy after mastectomy and axillary surgery on 10-year recurrence and 20-year breast cancer mortality: meta-analysis of individual patient data for 8135 women in 22 randomised trials.* Lancet, 2014.
136. Krag, D.N. and R.M. Single, *Breast cancer survival according to number of nodes removed.* Ann Surg Oncol, 2003. **10**(10): p. 1152-9.
137. Moon, H.G., W. Han, and D.Y. Noh, *Comparable survival between pN0 breast cancer patients undergoing sentinel node biopsy and extensive axillary dissection: a report from the Korean Breast Cancer Society.* J Clin Oncol, 2010. **28**(10): p. 1692-9.
138. Fisher, B., et al., *Twenty-five-year follow-up of a randomized trial comparing radical mastectomy, total mastectomy, and total mastectomy followed by irradiation.* N Engl J Med, 2002. **347**(8): p. 567-75.
139. Orr, R.K., *The impact of prophylactic axillary node dissection on breast cancer survival--a Bayesian meta-analysis.* Ann Surg Oncol, 1999. **6**(1): p. 109-16.

140. Krag, D.N., et al., *Sentinel-lymph-node resection compared with conventional axillary-lymph-node dissection in clinically node-negative patients with breast cancer: overall survival findings from the NSABP B-32 randomised phase 3 trial.* Lancet Oncol, 2010. **11**(10): p. 927-33.
141. Veronesi, U., et al., *Sentinel-lymph-node biopsy as a staging procedure in breast cancer: update of a randomised controlled study.* Lancet Oncol, 2006. **7**(12): p. 983-90.
142. Veronesi, U., et al., *A randomized comparison of sentinel-node biopsy with routine axillary dissection in breast cancer.* N Engl J Med, 2003. **349**(6): p. 546-53.
143. Giuliano, A.E., et al., *Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis: a randomized clinical trial.* JAMA, 2011. **305**(6): p. 569-75.
144. Singletary, S.E., et al., *A role for curative surgery in the treatment of selected patients with metastatic breast cancer.* Oncologist, 2003. **8**(3): p. 241-51.
145. Ruiterkamp, J., et al., *Impact of breast surgery on survival in patients with distant metastases at initial presentation: a systematic review of the literature.* Breast Cancer Res Treat, 2010. **120**(1): p. 9-16.
146. Badwe, R., *Surgical removal of primary tumor and axillary lymph nodes in women with metastatic breast cancer at first presentation: A randomized controlled trial*, in *San Antonio Breast Cancer Symposium*. 2013: San Antonio.
147. Early Breast Cancer Trialists' Collaborative, G., *Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials.* Lancet, 2005. **365**(9472): p. 1687-717.
148. Dowsett, M., et al., *Meta-analysis of breast cancer outcomes in adjuvant trials of aromatase inhibitors versus tamoxifen.* J Clin Oncol, 2010. **28**(3): p. 509-18.
149. Early Breast Cancer Trialists' Collaborative, G., et al., *Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials.* Lancet, 2012. **379**(9814): p. 432-44.
150. Dahabreh, I.J., et al., *Trastuzumab in the adjuvant treatment of early-stage breast cancer: a systematic review and meta-analysis of randomized controlled trials.* Oncologist, 2008. **13**(6): p. 620-30.
151. Mieog, J.S., J.A. van der Hage, and C.J. van de Velde, *Neoadjuvant chemotherapy for operable breast cancer.* Br J Surg, 2007. **94**(10): p. 1189-200.
152. Charehbili, A., et al., *Neoadjuvant hormonal therapy for endocrine sensitive breast cancer: a systematic review.* Cancer Treat Rev, 2014. **40**(1): p. 86-92.
153. Dent, S., et al., *HER2-targeted therapy in breast cancer: a systematic review of neoadjuvant trials.* Cancer Treat Rev, 2013. **39**(6): p. 622-31.
154. Hanrahan, E.O., B.T. Hennessy, and V. Valero, *Neoadjuvant systemic therapy for breast cancer: an overview and review of recent clinical trials.* Expert Opin Pharmacother, 2005. **6**(9): p. 1477-91.
155. Cortazar, P., et al., *Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis.* Lancet, 2014.

156. von Minckwitz, G., et al., *Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes.* J Clin Oncol, 2012. **30**(15): p. 1796-804.
157. Cardoso, F., et al., *International guidelines for management of metastatic breast cancer: combination vs sequential single-agent chemotherapy.* J Natl Cancer Inst, 2009. **101**(17): p. 1174-81.
158. Zhang, M., et al., *Chinese medicinal herbs to treat the side-effects of chemotherapy in breast cancer patients.* Cochrane Database Syst Rev, 2007(2): p. CD004921.
159. Azim, H.A., Jr., et al., *Long-term toxic effects of adjuvant chemotherapy in breast cancer.* Ann Oncol, 2011. **22**(9): p. 1939-47.
160. Amir, E., et al., *Toxicity of adjuvant endocrine therapy in postmenopausal breast cancer patients: a systematic review and meta-analysis.* J Natl Cancer Inst, 2011. **103**(17): p. 1299-309.
161. Hurria, A., S.E. Come, and L.J. Pierce. *Patterns of relapse and long-term complications of therapy in breast cancer survivors*. UptoDate 2014 10 Mar 2014 [cited 2014 27 June].
162. Emens, L.A. and N.E. Davidson, *The follow-up of breast cancer.* Semin Oncol, 2003. **30**(3): p. 338-48.
163. Borst, M.J. and J.A. Ingold, *Metastatic patterns of invasive lobular versus invasive ductal carcinoma of the breast.* Surgery, 1993. **114**(4): p. 637-41; discussion 641-2.
164. Bertos, N.R. and M. Park, *Breast cancer - one term, many entities?* J Clin Invest, 2011. **121**(10): p. 3789-96.
165. Saphner, T., D.C. Tormey, and R. Gray, *Annual hazard rates of recurrence for breast cancer after primary therapy.* J Clin Oncol, 1996. **14**(10): p. 2738-46.
166. Chen, Y., et al., *Epidemiology of contralateral breast cancer.* Cancer Epidemiol Biomarkers Prev, 1999. **8**(10): p. 855-61.
167. Mellemkjaer, L., et al., *Risk of second cancer among women with breast cancer.* Int J Cancer, 2006. **118**(9): p. 2285-92.
168. Falk, R.S., et al., *Second events following ductal carcinoma in situ of the breast: a register-based cohort study.* Breast Cancer Res Treat, 2011. **129**(3): p. 929-38.
169. Innos, K. and P.L. Horn-Ross, *Risk of second primary breast cancers among women with ductal carcinoma in situ of the breast.* Breast Cancer Res Treat, 2008. **111**(3): p. 531-40.
170. Lizarraga, I.M., et al., *Review of risk factors for the development of contralateral breast cancer.* Am J Surg, 2013. **206**(5): p. 704-8.
171. Hartman, M., et al., *Incidence and prognosis of synchronous and metachronous bilateral breast cancer.* J Clin Oncol, 2007. **25**(27): p. 4210-6.
172. Matesich, S.M. and C.L. Shapiro, *Second cancers after breast cancer treatment.* Semin Oncol, 2003. **30**(6): p. 740-8.
173. Prochazka, M., et al., *Family history of breast cancer and young age at diagnosis of breast cancer increase risk of second primary malignancies in women: a population-based cohort study.* Br J Cancer, 2006. **95**(9): p. 1291-5.
174. *Cancer Survival in Singapore 1968-2007*. 2011, Singapore Cancer Registry.

175. Coleman, M.P., et al., *Cancer survival in five continents: a worldwide population-based study (CONCORD).* Lancet Oncol, 2008. **9**(8): p. 730-56.
176. Rachet, B., et al., *Population-based cancer survival trends in England and Wales up to 2007: an assessment of the NHS cancer plan for England.* Lancet Oncol, 2009. **10**(4): p. 351-69.
177. Coleman, M.P., P. Babb, and P. Damiecki, *Cancer survival trends in England and Wales 1971-1995: deprivation and NHS region*, in *SMPS*. 1999, Stationery Office: London.
178. Bhoo-Pathy, N., et al., *Ethnic differences in survival after breast cancer in South East Asia.* PLoS One, 2012. **7**(2): p. e30995.
179. Yu, X.Q., *Socioeconomic disparities in breast cancer survival: relation to stage at diagnosis, treatment and race.* BMC Cancer, 2009. **9**: p. 364.
180. Harper, S., et al., *Trends in area-socioeconomic and race-ethnic disparities in breast cancer incidence, stage at diagnosis, screening, mortality, and survival among women ages 50 years and over (1987-2005).* Cancer Epidemiol Biomarkers Prev, 2009. **18**(1): p. 121-31.
181. Newman, L.A., et al., *Meta-analysis of survival in African American and white American patients with breast cancer: ethnicity compared with socioeconomic status.* J Clin Oncol, 2006. **24**(9): p. 1342-9.
182. Gomez, S.L., et al., *Disparities in breast cancer survival among Asian women by ethnicity and immigrant status: a population-based study.* Am J Public Health, 2010. **100**(5): p. 861-9.
183. Maskarinec, G., et al., *Ethnic differences in breast cancer survival: status and determinants.* Womens Health (Lond Engl), 2011. **7**(6): p. 677-87.
184. Kogevinas, M. and M. Porta, *Socioeconomic differences in cancer survival: a review of the evidence.* IARC Sci Publ, 1997(138): p. 177-206.
185. Engelhardt, E.G., et al., *Predicting and communicating the risk of recurrence and death in women with early-stage breast cancer: a systematic review of risk prediction models.* J Clin Oncol, 2014. **32**(3): p. 238-50.
186. Haybittle, J.L., et al., *A prognostic index in primary breast cancer.* Br J Cancer, 1982. **45**(3): p. 361-6.
187. Todd, J.H., et al., *Confirmation of a prognostic index in primary breast cancer.* Br J Cancer, 1987. **56**(4): p. 489-92.
188. Ravdin, P.M., et al., *Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer.* J Clin Oncol, 2001. **19**(4): p. 980-91.
189. Michaelson, J.S., et al., *Improved web-based calculators for predicting breast carcinoma outcomes.* Breast Cancer Res Treat, 2011. **128**(3): p. 827-35.
190. Hornberger, J., et al., *Clinical validity/utility, change in practice patterns, and economic implications of risk stratifiers to predict outcomes for early-stage breast cancer: a systematic review.* J Natl Cancer Inst, 2012. **104**(14): p. 1068-79.
191. Harnett, A., et al., *Diagnosis and treatment of early breast cancer, including locally advanced disease--summary of NICE guidance.* BMJ, 2009. **338**: p. b438.

192. Agarwal, V. and P. O'Neill, *Adjuvant! Online as a decision-making tool in early breast cancer--a UK national survey.* Clin Oncol (R Coll Radiol), 2011. **23**(2): p. 159-60.
193. Mook, S., et al., *Calibration and discriminatory accuracy of prognosis calculation for breast cancer with the online Adjuvant! program: a hospital-based retrospective cohort study.* Lancet Oncol, 2009. **10**(11): p. 1070-6.
194. Epstein, R.J., et al., *Utility of a web-based breast cancer predictive algorithm for adjuvant chemotherapeutic decision making in a multidisciplinary oncology center.* Cancer Invest, 2006. **24**(4): p. 367-73.
195. Campbell, H.E., et al., *An investigation into the performance of the Adjuvant! Online prognostic programme in early breast cancer for a cohort of patients in the United Kingdom.* Br J Cancer, 2009. **101**(7): p. 1074-84.
196. Olivotto, I.A., et al., *Population-based validation of the prognostic model ADJUVANT! for early breast cancer.* J Clin Oncol, 2005. **23**(12): p. 2716-25.
197. Bhoo-Pathy, N., et al., *Adjuvant! Online is overoptimistic in predicting survival of Asian breast cancer patients.* Eur J Cancer, 2012. **48**(7): p. 982-9.
198. Jung, M., et al., *Application of the adjuvant! Online model to Korean breast cancer patients: an assessment of prognostic accuracy and development of an alternative prognostic tool.* Ann Surg Oncol, 2013. **20**(8): p. 2615-24.
199. Yao-Lung, K., C. Dar-Ren, and C. Tsai-Wang, *Accuracy validation of adjuvant! online in Taiwanese breast cancer patients--a 10-year analysis.* BMC Med Inform Decis Mak, 2012. **12**: p. 108.
200. Agarwal, G., et al., *Spectrum of breast cancer in Asian women.* World J Surg, 2007. **31**(5): p. 1031-40.
201. Porter, P., *"Westernizing" women's risks? Breast cancer in lower-income countries.* N Engl J Med, 2008. **358**(3): p. 213-6.
202. Kamangar, F., G.M. Dores, and W.F. Anderson, *Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world.* J Clin Oncol, 2006. **24**(14): p. 2137-50.
203. Leggett, L.E., et al., *Experiences and attitudes toward risk of recurrence testing in women with breast cancer: a systematic review.* Breast Cancer Res Treat, 2014. **144**(3): p. 457-65.
204. O'Neill, S.C., et al., *Women's interest in gene expression analysis for breast cancer recurrence risk.* J Clin Oncol, 2007. **25**(29): p. 4628-34.
205. Kattan, M.W., et al., *A tool for predicting breast carcinoma mortality in women who do not receive adjuvant therapy.* Cancer, 2004. **101**(11): p. 2509-15.
206. Campbell, H.E., et al., *Estimation and external validation of a new prognostic model for predicting recurrence-free survival for early breast cancer patients in the UK.* Br J Cancer, 2010. **103**(6): p. 776-86.
207. Wishart, G.C., et al., *A population-based validation of the prognostic model PREDICT for early breast cancer.* Eur J Surg Oncol, 2011. **37**(5): p. 411-7.

208. Wishart, G.C., et al., *PREDICT Plus: development and validation of a prognostic model for early breast cancer that includes HER2.* Br J Cancer, 2012. **107**(5): p. 800-7.
209. Sigurdsson, H., et al., *Indicators of prognosis in node-negative breast cancer.* N Engl J Med, 1990. **322**(15): p. 1045-53.
210. Aubele, M., et al., *Disease-free survival of node-positive breast cancer patients. Improved prognostication by cytometrical parameters.* Pathol Res Pract, 1995. **191**(10): p. 982-90.
211. Chang, H.Y., et al., *Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds.* PLoS Biol, 2004. **2**(2): p. E7.
212. Pawitan, Y., et al., *Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts.* Breast Cancer Res, 2005. **7**(6): p. R953-64.
213. Wang, Y., et al., *Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer.* Lancet, 2005. **365**(9460): p. 671-9.
214. Ring, B.Z., et al., *Novel prognostic immunohistochemical biomarker panel for estrogen receptor-positive breast cancer.* J Clin Oncol, 2006. **24**(19): p. 3039-47.
215. Ma, Y., et al., *Population-based molecular prognosis of breast cancer by transcriptional profiling.* Clin Cancer Res, 2007. **13**(7): p. 2014-22.
216. Xu, L., et al., *Merging microarray data from separate breast cancer studies provides a robust prognostic test.* BMC Bioinformatics, 2008. **9**: p. 125.
217. Ma, X.J., et al., *A five-gene molecular grade index and HOXB13:IL17BR are complementary prognostic factors in early stage breast cancer.* Clin Cancer Res, 2008. **14**(9): p. 2601-8.
218. Parker, J.S., et al., *Supervised risk predictor of breast cancer based on intrinsic subtypes.* J Clin Oncol, 2009. **27**(8): p. 1160-7.
219. Haibe-Kains, B., et al., *A fuzzy gene expression-based computational approach improves breast cancer prognostication.* Genome Biol, 2010. **11**(2): p. R18.
220. Van Laar, R.K., *Design and multiseries validation of a web-based gene expression assay for predicting breast cancer recurrence and patient survival.* J Mol Diagn, 2011. **13**(3): p. 297-304.
221. Wang, J.J. and J. Attia, *Study designs in epidemiology and levels of evidence.* Am J Ophthalmol, 2010. **149**(3): p. 367-70.
222. Grobbee, D.E. and A.W. Hoes, *Clinical Epidemiology: Principles, Methods, and Applications for Clinical Research* 1ed. 2008: Jones and Bartlett Publisers. 413.
223. Moons, K.G., et al., *Prognosis and prognostic research: what, why, and how?* BMJ, 2009. **338**: p. b375.
224. Moons, K.G., et al., *Prognosis and prognostic research: application and impact of prognostic models in clinical practice.* BMJ, 2009. **338**: p. b606.
225. Burns, P.B., R.J. Rohrich, and K.C. Chung, *The levels of evidence and their role in evidence-based medicine.* Plast Reconstr Surg, 2011. **128**(1): p. 305-10.
226. Song, J.W. and K.C. Chung, *Observational studies: cohort and case-control studies.* Plast Reconstr Surg, 2010. **126**(6): p. 2234-42.

227. Realini, J.P., et al., *Oral contraceptives and venous thromboembolism: a case-control study designed to minimize detection bias.* J Am Board Fam Pract, 1997. **10**(5): p. 315-21.
228. Chasan-Taber, L., et al., *Prospective study of oral contraceptives and hypertension among women in the United States.* Circulation, 1996. **94**(3): p. 483-9.
229. Lewallen, S. and P. Courtright, *Epidemiology in practice: case-control studies.* Community Eye Health, 1998. **11**(28): p. 57-8.
230. Levin, K.A., *Study design III: Cross-sectional studies.* Evid Based Dent, 2006. **7**(1): p. 24-5.
231. Gjerstorff, M.L., *The Danish Cancer Registry.* Scand J Public Health, 2011. **39**(7 Suppl): p. 42-5.
232. Storm, H.H., et al., *The Danish Cancer Registry--history, content, quality and use.* Dan Med Bull, 1997. **44**(5): p. 535-9.
233. Pukkala, E., et al., *Cancer atlas of Northern Europe*. 2001, Cancer Society of Finland: Helsinki
234. Ording, A.G., et al., *Completeness of breast cancer staging in the Danish Cancer Registry, 2004-2009.* Clin Epidemiol, 2012. **4 Suppl 2**: p. 11-6.
235. Teppo, L., E. Pukkala, and M. Lehtonen, *Data quality and quality control of a population-based cancer registry. Experience in Finland.* Acta Oncol, 1994. **33**(4): p. 365-9.
236. *Cancer in Finland 2006 and 2007*. 2009, Finnish Cancer Registry.
237. Herrmann, C., et al., *Cancer survivors in Switzerland: a rapidly growing population to care for.* BMC Cancer, 2013. **13**: p. 287.
238. Verkooijen, H.M., et al., *Set-up of a population-based familial breast cancer registry in Geneva, Switzerland: validation of first results.* Ann Oncol, 2004. **15**(2): p. 350-3.
239. Bouchardy, C., E. Rapiti, and S. Benhamou, *Cancer registries can provide evidence-based data to improve quality of care and prevent cancer deaths.* Ecancermedicalscience, 2014. **8**: p. 413.
240. Taban, F., et al., *Breast cancer management and outcome according to surgeon's affiliation: a population-based comparison adjusted for patient's selection bias.* Ann Oncol, 2013. **24**(1): p. 116-25.
241. *Cancer in Norway 2011 - Cancer incidence, mortality, survival and prevalence in Norway*, I.K. Larsen, Editor. 2013, Cancer Registry of Norway: Oslo.
242. Larsen, I.K., et al., *Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness.* Eur J Cancer, 2009. **45**(7): p. 1218-31.
243. Chia, K.S., et al., *Trends in Cancer Incidence in Singapore 1993-1997*. 2000, Singapore Cancer Registry.
244. Seow, A., et al., *Trends in Cancer Incidence in Singapore 1968-2002*. 2004, Singapore Cancer Registry
245. Barlow, L., et al., *The completeness of the Swedish Cancer Register: a sample survey for year 1998.* Acta Oncol, 2009. **48**(1): p. 27-33.
246. Mattsson, B. and A. Wallgren, *Completeness of the Swedish Cancer Register. Non-notified cancer cases recorded on death certificates in 1978.* Acta Radiol Oncol, 1984. **23**(5): p. 305-13.
247. *Multi-generation register 2009: A description of contents and quality*, in *Background Facts*. Statistics Sweden.

248. Ekbom, A., *The Swedish Multi-generation Register.* Methods Mol Biol, 2011. **675**: p. 215-20.
249. Warnberg, F., J. Bergh, and L. Holmberg, *Prognosis in women with a carcinoma in situ of the breast: a population-based study in Sweden.* Cancer Epidemiol Biomarkers Prev, 1999. **8**(9): p. 769-74.
250. Garne, J.P., K. Aspegren, and T. Moller, *Validity of breast cancer registration from one hospital into the Swedish National Cancer Registry 1971-1991.* Acta Oncol, 1995. **34**(2): p. 153-6.
251. Khan, K.S., et al., *Five steps to conducting a systematic review.* J R Soc Med, 2003. **96**(3): p. 118-21.
252. Moher, D., et al., *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.* BMJ, 2009. **339**: p. b2535.
253. Moons, K.G., et al., *Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist.* PLoS Med, 2014. **11**(10): p. e1001744.
254. Kleinbaum, D.G. and M. Klein, *Survival Analysis : A self-learning Text*. 2nd ed. Statistics for Biology and Health, ed. M. Gail, et al. 2005: Springer.
255. Parkin, D.M. and T. Hakulinen, *Analysis of survival*, in *Cancer Registration: Principles and Methods*, O.M. Jensen, et al., Editors. 1991, International Agency for Research on Cancer: Lyon, France. p. 164.
256. Dickman, P.W., et al., *Regression models for relative survival.* Stat Med, 2004. **23**(1): p. 51-64.
257. Ederer, F., L.M. Axtell, and S.J. Cutler, *The relative survival rate: a statistical methodology.* Natl Cancer Inst Monogr, 1961. **6**: p. 101-21.
258. Ederer, F. and H. Heise, *Instructions to IBM 650 programmers in processing survival computations. Methodological note no.10* 1959, National Cancer Insititute
259. Hakulinen, T., *Cancer survival corrected for heterogeneity in patient withdrawal.* Biometrics, 1982. **38**(4): p. 933-42.
260. Sahai, H. and A. Khurshid, *Confidence-Intervals for the Mean of a Poisson-Distribution - a Review.* Biometrical Journal, 1993. **35**(7): p. 857-867.
261. Breslow, N.E. and N.E. Day, *Statistical methods in Cancer Research Volume II- The Design and Analysis of Cohort Studies*. 1987: International Agency For Research on Cancer.
262. Christensen, E., *Multivariate survival analysis using Cox's regression model.* Hepatology, 1987. **7**(6): p. 1346-58.
263. Tai, B.C. and D. Machin, *Regression Methods for Medical Research*. 2013: Wiley Blackwell.
264. Royston, P. and D.G. Altman, *External validation of a Cox prognostic model: principles and methods.* BMC Med Res Methodol, 2013. **13**: p. 33.
265. Viel, J.F., *[Poisson regression in epidemiology]*. Rev Epidemiol Sante Publique, 1994. **42**(1): p. 79-87.
266. Carstensen, B. *Demography and epidemiology: Practical use of the Lexis diagram in the computer age or "Who needs the Cox model anyway?"*. in *Annual meeting of Finnish Statistical Society*. 2005. Oulu.
267. Collins, G.S., et al., *External validation of multivariable prediction models: a systematic review of methodological conduct and reporting.* BMC Med Res Methodol, 2014. **14**: p. 40.

268. Steyerberg, E.W., et al., *Prognosis Research Strategy (PROGRESS) 3: prognostic model research.* PLoS Med, 2013. **10**(2): p. e1001381.
269. Bleeker, S.E., et al., *External validation is necessary in prediction research: a clinical example.* J Clin Epidemiol, 2003. **56**(9): p. 826-32.
270. Smith, G.C., et al., *Correcting for optimistic prediction in small data sets.* Am J Epidemiol, 2014. **180**(3): p. 318-24.
271. May, M., et al., *Development and validation of a prognostic model for survival time data: application to prognosis of HIV positive patients treated with antiretroviral therapy.* Stat Med, 2004. **23**(15): p. 2375-98.
272. Harrell, F.E., Jr., K.L. Lee, and D.B. Mark, *Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors.* Stat Med, 1996. **15**(4): p. 361-87.
273. Debray, T.P., et al., *A new framework to enhance the interpretation of external validation studies of clinical prediction models.* J Clin Epidemiol, 2014.
274. Goss, P.E., et al., *Male breast carcinoma: a review of 229 patients who presented to the Princess Margaret Hospital during 40 years: 1955-1996.* Cancer, 1999. **85**(3): p. 629-39.
275. Jemal, A., et al., *Cancer statistics, 2004.* CA Cancer J Clin, 2004. **54**(1): p. 8-29.
276. Czene, K., et al., *How to treat male breast cancer.* Breast, 2007. **16 Suppl 2**: p. S147-54.
277. Anderson, W.F., et al., *Is male breast cancer similar or different than female breast cancer?* Breast Cancer Res Treat, 2004. **83**(1): p. 77-86.
278. Ravandi-Kashani, F. and T.G. Hayes, *Male breast cancer: a review of the literature.* Eur J Cancer, 1998. **34**(9): p. 1341-7.
279. Giordano, S.H., *A review of the diagnosis and management of male breast cancer.* Oncologist, 2005. **10**(7): p. 471-9.
280. El-Tamer, M.B., et al., *Men with breast cancer have better disease-specific survival than women.* Arch Surg, 2004. **139**(10): p. 1079-82.
281. Kalager, M., et al., *Effect of screening mammography on breast-cancer mortality in Norway.* N Engl J Med, 2010. **363**(13): p. 1203-10.
282. Scheike, O., *Male breast cancer. 5. Clinical manifestations in 257 cases in Denmark.* Br J Cancer, 1973. **28**(6): p. 552-61.
283. Anelli, T.F., et al., *Tamoxifen administration is associated with a high rate of treatment-limiting symptoms in male breast cancer patients.* Cancer, 1994. **74**(1): p. 74-7.
284. Gennari, R., et al., *Male breast cancer: a special therapeutic problem. Anything new? (Review).* Int J Oncol, 2004. **24**(3): p. 663-70.
285. Visram, H., F. Kanji, and S.F. Dent, *Endocrine therapy for male breast cancer: rates of toxicity and adherence.* Curr Oncol, 2010. **17**(5): p. 17-21.
286. Harlan, L.C., et al., *Breast cancer in men in the United States: a population-based study of diagnosis, treatment, and survival.* Cancer, 2010. **116**(15): p. 3558-68.
287. Heller, K.S., et al., *Male breast cancer: a clinicopathologic study of 97 cases.* Ann Surg, 1978. **188**(1): p. 60-5.
288. Chuba, P.J., et al., *Bilateral risk for subsequent breast cancer after lobular carcinoma-in-situ: analysis of surveillance, epidemiology, and end results data.* J Clin Oncol, 2005. **23**(24): p. 5534-41.

289. Erbas, B., et al., *The natural history of ductal carcinoma in situ of the breast: a review.* Breast Cancer Res Treat, 2006. **97**(2): p. 135-44.
290. Franceschi, S., et al., *Second cancers following in situ carcinoma of the breast.* Int J Cancer, 1998. **77**(3): p. 392-5.
291. Li, C.I., et al., *Risk of invasive breast carcinoma among women diagnosed with ductal carcinoma in situ and lobular carcinoma in situ, 1988-2001.* Cancer, 2006. **106**(10): p. 2104-12.
292. Robinson, D., L. Holmberg, and H. Moller, *The occurrence of invasive cancers following a diagnosis of breast carcinoma in situ.* Br J Cancer, 2008. **99**(4): p. 611-5.
293. Habel, L.A., et al., *Risk of contralateral breast cancer among women with carcinoma in situ of the breast.* Ann Surg, 1997. **225**(1): p. 69-75.
294. Rawal, R., J. Lorenzo Bermejo, and K. Hemminki, *Risk of subsequent invasive breast carcinoma after in situ breast carcinoma in a population covered by national mammographic screening.* Br J Cancer, 2005. **92**(1): p. 162-6.
295. Levi, F., et al., *Invasive breast cancer following ductal and lobular carcinoma in situ of the breast.* Int J Cancer, 2005. **116**(5): p. 820-3.
296. Warnberg, F., J. Yuen, and L. Holmberg, *Risk of subsequent invasive breast cancer after breast carcinoma in situ.* Lancet, 2000. **355**(9205): p. 724-5.
297. Soerjomataram, I., et al., *Increased risk of second malignancies after in situ breast carcinoma in a population-based registry.* Br J Cancer, 2006. **95**(3): p. 393-7.
298. Eusebi, V., et al., *Long-term follow-up of in situ carcinoma of the breast.* Semin Diagn Pathol, 1994. **11**(3): p. 223-35.
299. Hemminki, K. and C. Granstrom, *Familial breast carcinoma risks by morphology: a nationwide epidemiologic study from Sweden.* Cancer, 2002. **94**(11): p. 3063-70.
300. Warnberg, F., et al., *Risk factors for subsequent invasive breast cancer and breast cancer death after ductal carcinoma in situ: a population-based case-control study in Sweden.* Cancer Epidemiol Biomarkers Prev, 2001. **10**(5): p. 495-9.
301. Emdin, S.O., et al., *SweDCIS: Radiotherapy after sector resection for ductal carcinoma in situ of the breast. Results of a randomised trial in a population offered mammography screening.* Acta Oncol, 2006. **45**(5): p. 536-43.
302. Claus, E.B., et al., *Prevalence of BRCA1 and BRCA2 mutations in women diagnosed with ductal carcinoma in situ.* JAMA, 2005. **293**(8): p. 964-9.
303. Harris, E.E., et al., *Relationship of family history and outcome after breast conservation therapy in women with ductal carcinoma in situ of the breast.* Int J Radiat Oncol Biol Phys, 2000. **48**(4): p. 933-41.
304. Ji, J. and K. Hemminki, *Risk for contralateral breast cancers in a population covered by mammography: effects of family history, age at diagnosis and histology.* Breast Cancer Res Treat, 2007. **105**(2): p. 229-36.
305. Bradley, S.J., D.W. Weaver, and D.L. Bouwman, *Alternatives in the surgical management of in situ breast cancer. A meta-analysis of outcome.* Am Surg, 1990. **56**(7): p. 428-32.
306. Collaborative Group on Hormonal Factors in Breast, C., *Familial breast cancer: collaborative reanalysis of individual data from 52*

epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. Lancet, 2001. **358**(9291): p. 1389-99.
307. Olsson, S., et al., *Implementation of service screening with mammography in Sweden: from pilot study to nationwide programme.* J Med Screen, 2000. **7**(1): p. 14-8.
308. *Breast Cancer. Report from the Regional Cancer Register Stockholm-Gotland Region 1977-2006*. 2011, ONKOLOGISKT CENTRUM STOCKHOLM GOTLAND.
309. Clarke, M., *Meta-analyses of adjuvant therapies for women with early breast cancer: the Early Breast Cancer Trialists' Collaborative Group overview.* Ann Oncol, 2006. **17 Suppl 10**: p. x59-62.
310. Early Breast Cancer Trialists' Collaborative, G., et al., *Adjuvant chemotherapy in oestrogen-receptor-poor breast cancer: patient-level meta-analysis of randomised trials.* Lancet, 2008. **371**(9606): p. 29-40.
311. Higgins, M.J. and J. Baselga, *Targeted therapies for breast cancer.* J Clin Invest, 2011. **121**(10): p. 3797-803.
312. Liddell, F.D., *Simple exact analysis of the standardised mortality ratio.* J Epidemiol Community Health, 1984. **38**(1): p. 85-8.
313. Chen, L.L., et al., *Technical Repor #10 -Technical report for the paper: "CancerMath.net: Wed-based Calculators for Breast Carcinoma" (July 28, 2009, updated November 28, 2009 and March 22, 2010) at http://cancer.lifemath.net/about/techreports/index.php* 2010.
314. Tan, B.K., et al., *Do Asian breast cancer patients have poorer survival than their western counterparts? A comparison between Singapore and Stockholm.* Breast Cancer Res, 2009. **11**(1): p. R4.
315. Lim, S.E., et al., *Clinical observations from a breast cancer registry in Asian women.* World J Surg, 2007. **31**(7): p. 1387-92.
316. Saxena, N., et al., *Breast cancer in South East Asia: comparison of presentation and outcome between a middle income and a high income country.* World J Surg, 2012. **36**(12): p. 2838-46.
317. Prehn, A.W., et al., *Differences in treatment patterns for localized breast carcinoma among Asian/Pacific islander women.* Cancer, 2002. **95**(11): p. 2268-75.
318. Iskandarsyah, A., et al., *Psychosocial and cultural reasons for delay in seeking help and nonadherence to treatment in Indonesian women with breast cancer: a qualitative study.* Health Psychol, 2014. **33**(3): p. 214-21.
319. Lee, M.M., et al., *Alternative therapies used by women with breast cancer in four ethnic populations.* J Natl Cancer Inst, 2000. **92**(1): p. 42-7.
320. Leong, B.D., et al., *Trends of breast cancer treatment in Sabah, Malaysia: a problem with lack of awareness.* Singapore Med J, 2009. **50**(8): p. 772-6.
321. Chong, H.Y., et al., *Treatment options for locally advanced breast cancer--experience in an Asian tertiary hospital.* Asian Pac J Cancer Prev, 2010. **11**(4): p. 913-7.
322. Kardinah, D., et al., *Short report: Limited effectiveness of screening mammography in addition to clinical breast examination by trained nurse midwives in rural Jakarta, Indonesia.* Int J Cancer, 2014. **134**(5): p. 1250-5.
323. Phan, V.H., et al., *Ethnic differences in drug metabolism and toxicity from chemotherapy.* Expert Opin Drug Metab Toxicol, 2009. **5**(3): p. 243-57.

324. Na, K.Y., et al., *The 70-gene prognostic signature for korean breast cancer patients.* J Breast Cancer, 2011. **14**(1): p. 33-8.
325. Ishitobi, M., et al., *Clinical utility of the 70-gene MammaPrint profile in a Japanese population.* Jpn J Clin Oncol, 2010. **40**(6): p. 508-12.
326. Yip, C.H., N.A. Taib, and I. Mohamed, *Epidemiology of breast cancer in Malaysia.* Asian Pac J Cancer Prev, 2006. **7**(3): p. 369-74.
327. Tan, E.Y., et al., *Locally advanced and metastatic breast cancer in a tertiary hospital.* Ann Acad Med Singapore, 2005. **34**(10): p. 595-601.
328. Chopra, R., *The Indian scene.* J Clin Oncol, 2001. **19**(18 Suppl): p. 106S-111S.
329. Sant, M., et al., *Breast carcinoma survival in Europe and the United States.* Cancer, 2004. **100**(4): p. 715-22.
330. Ly, B.H., et al., *Loco-regional treatment in metastatic breast cancer patients: is there a survival benefit?* Breast Cancer Res Treat, 2010. **119**(3): p. 537-45.
331. Chang, J., et al., *Survival of patients with metastatic breast carcinoma: importance of prognostic markers of the primary tumor.* Cancer, 2003. **97**(3): p. 545-53.
332. Khodari, W., et al., *Impact of loco-regional treatment on metastatic breast cancer outcome: A review.* Crit Rev Oncol Hematol, 2013.
333. Perez-Fidalgo, J.A., et al., *Removal of primary tumor improves survival in metastatic breast cancer. Does timing of surgery influence outcomes?* Breast, 2011. **20**(6): p. 548-54.
334. Tai, P., et al., *Survival of patients with metastatic breast cancer: twenty-year data from two SEER registries.* BMC Cancer, 2004. **4**: p. 60.
335. Todd, M., M. Shoag, and E. Cadman, *Survival of women with metastatic breast cancer at Yale from 1920 to 1980.* J Clin Oncol, 1983. **1**(6): p. 406-8.
336. Largillier, R., et al., *Prognostic factors in 1,038 women with metastatic breast cancer.* Ann Oncol, 2008. **19**(12): p. 2012-9.
337. Andre, F., et al., *Breast cancer with synchronous metastases: trends in survival during a 14-year period.* J Clin Oncol, 2004. **22**(16): p. 3302-8.
338. Rapiti, E., et al., *Complete excision of primary breast tumor improves survival of patients with metastatic breast cancer at diagnosis.* J Clin Oncol, 2006. **24**(18): p. 2743-9.
339. Pathy, N.B., et al., *Impact of breast surgery on survival in women presenting with metastatic breast cancer.* Br J Surg, 2011. **98**(11): p. 1566-72.
340. Degner, L.F., et al., *Information needs and decisional preferences in women with breast cancer.* JAMA, 1997. **277**(18): p. 1485-92.
341. Glare, P., et al., *A systematic review of physicians' survival predictions in terminally ill cancer patients.* BMJ, 2003. **327**(7408): p. 195-8.
342. Moher, D., et al., *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.* PLoS Med, 2009. **6**(7): p. e1000097.
343. von Elm, E., et al., *The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies.* Lancet, 2007. **370**(9596): p. 1453-7.
344. Donders, A.R., et al., *Review: a gentle introduction to imputation of missing values.* J Clin Epidemiol, 2006. **59**(10): p. 1087-91.

345. Gordon, L. and R.A. Olshen, *Tree-structured survival analysis.* Cancer Treat Rep, 1985. **69**(10): p. 1065-9.
346. Biganzoli, E., et al., *Feed forward neural networks for the analysis of censored survival data: a partial logistic regression approach.* Stat Med, 1998. **17**(10): p. 1169-86.
347. Nash, C.H., 3rd, et al., *Prediction of outcome in metastatic breast cancer treated with adriamycin combination chemotherapy.* Cancer, 1980. **46**(11): p. 2380-8.
348. Hortobagyi, G.N., et al., *Multivariate analysis of prognostic factors in metastatic breast cancer.* J Clin Oncol, 1983. **1**(12): p. 776-86.
349. Williams, M.R., et al., *Survival patterns in hormone treated advanced breast cancer.* Br J Surg, 1986. **73**(9): p. 752-5.
350. Rabinovich, M., et al., *Development and validation of prognostic models in metastatic breast cancer: a GOCS study.* Oncology, 1992. **49**(3): p. 188-95.
351. Yamamoto, N., et al., *Construction and validation of a practical prognostic index for patients with metastatic breast cancer.* J Clin Oncol, 1998. **16**(7): p. 2401-8.
352. Ryberg, M., et al., *Prognostic factors and long-term survival in 585 patients with metastatic breast cancer treated with epirubicin-based chemotherapy.* Ann Oncol, 2001. **12**(1): p. 81-7.
353. Giordano, A., et al., *Artificial neural network analysis of circulating tumor cells in metastatic breast cancer patients.* Breast Cancer Res Treat, 2011. **129**(2): p. 451-8.
354. Giordano, A., et al., *Establishment and validation of circulating tumor cell-based prognostic nomograms in first-line metastatic breast cancer patients.* Clin Cancer Res, 2013. **19**(6): p. 1596-602.
355. Le Scodan, R., et al., *Brain metastases from breast carcinoma: validation of the radiation therapy oncology group recursive partitioning analysis classification and proposition of a new prognostic score.* Int J Radiat Oncol Biol Phys, 2007. **69**(3): p. 839-45.
356. Nieder, C., et al., *Prognostic scores in brain metastases from breast cancer.* BMC Cancer, 2009. **9**: p. 105.
357. Sperduto, P.W., et al., *Effect of tumor subtype on survival and the graded prognostic assessment for patients with breast cancer and brain metastases.* Int J Radiat Oncol Biol Phys, 2012. **82**(5): p. 2111-7.
358. Ahn, H.K., et al., *Prediction of outcomes for patients with brain parenchymal metastases from breast cancer (BC): a new BC-specific prognostic model and a nomogram.* Neuro Oncol, 2012. **14**(8): p. 1105-13.
359. Marko, N.F., et al., *Predicting survival in women with breast cancer and brain metastasis: a nomogram outperforms current survival prediction models.* Cancer, 2012. **118**(15): p. 3749-57.
360. Le Scodan, R., et al., *Brain metastases from breast cancer: proposition of new prognostic score including molecular subtypes and treatment.* J Neurooncol, 2012. **106**(1): p. 169-76.
361. Niwinska, A. and M. Murawska, *New breast cancer recursive partitioning analysis prognostic index in patients with newly diagnosed brain metastases.* Int J Radiat Oncol Biol Phys, 2012. **82**(5): p. 2065-71.

362. Rades, D., S. Douglas, and S.E. Schild, *A validated survival score for breast cancer patients with metastatic spinal cord compression.* Strahlenther Onkol, 2013. **189**(1): p. 41-6.
363. Liu, M.T., et al., *Prediction of outcome of patients with metastatic breast cancer: evaluation with prognostic factors and Nottingham prognostic index.* Support Care Cancer, 2010. **18**(12): p. 1553-64.
364. Robertson, J.F., et al., *Confirmation of a prognostic index for patients with metastatic breast cancer treated by endocrine therapy.* Breast Cancer Res Treat, 1992. **22**(3): p. 221-7.
365. Yip, C.H., *Breast cancer in Asia.* Methods Mol Biol, 2009. **471**: p. 51-64.
366. Li, C.I., K.E. Malone, and J.R. Daling, *Differences in breast cancer hormone receptor status and histology by race and ethnicity among women 50 years of age and older.* Cancer Epidemiol Biomarkers Prev, 2002. **11**(7): p. 601-7.
367. Pal, S.K., et al., *Impact of modern chemotherapy on the survival of women presenting with de novo metastatic breast cancer.* BMC Cancer, 2012. **12**: p. 435.
368. Chia, S.K., et al., *The impact of new chemotherapeutic and hormone agents on survival in a population-based cohort of women with metastatic breast cancer.* Cancer, 2007. **110**(5): p. 973-9.
369. Gradishar, W.J., *Emerging approaches for treating HER2-positive metastatic breast cancer beyond trastuzumab.* Ann Oncol, 2013.
370. O'Shaughnessy, J., *Extending survival with chemotherapy in metastatic breast cancer.* Oncologist, 2005. **10 Suppl 3**: p. 20-9.
371. Cristofanilli, M., et al., *Circulating tumor cells, disease progression, and survival in metastatic breast cancer.* N Engl J Med, 2004. **351**(8): p. 781-91.
372. Cristofanilli, M., et al., *Circulating tumor cells: a novel prognostic factor for newly diagnosed metastatic breast cancer.* J Clin Oncol, 2005. **23**(7): p. 1420-30.
373. Balic, M., et al., *Comparison of two methods for enumerating circulating tumor cells in carcinoma patients.* Cytometry B Clin Cytom, 2005. **68**(1): p. 25-30.
374. Muller, V., et al., *Prognostic impact of circulating tumor cells assessed with the CellSearch System and AdnaTest Breast in metastatic breast cancer patients: the DETECT study.* Breast Cancer Res, 2012. **14**(4): p. R118.
375. Pemmaraju, N., et al., *Retrospective review of male breast cancer patients: analysis of tamoxifen-related side-effects.* Ann Oncol, 2012. **23**(6): p. 1471-4.
376. Korde, L.A., et al., *Multidisciplinary meeting on male breast cancer: summary and research recommendations.* J Clin Oncol, 2010. **28**(12): p. 2114-22.
377. Tan, P.H., *Pathology of ductal carcinoma in situ of the breast: a heterogeneous entity in need of greater understanding.* Ann Acad Med Singapore, 2001. **30**(6): p. 671-6; quiz 677.
378. Liu, Y., et al., *Accuracy of perceived risk of recurrence among patients with early-stage breast cancer.* Cancer Epidemiol Biomarkers Prev, 2010. **19**(3): p. 675-80.

379. Lopez, M.E., et al., *Ductal carcinoma in situ (DCIS): posttreatment follow-up care among Latina and non-Latina White women.* J Cancer Surviv, 2013. **7**(2): p. 219-26.
380. Welch, H.G., S. Woloshin, and L.M. Schwartz, *The sea of uncertainty surrounding ductal carcinoma in situ--the price of screening mammography.* J Natl Cancer Inst, 2008. **100**(4): p. 228-9.
381. Meyerson, A.F., et al., *Outcome of long term active surveillance for estrogen receptor-positive ductal carcinoma in situ.* Breast, 2011. **20**(6): p. 529-33.
382. Wang, S.C., *The Singapore National Breast Screening Programme: principles and implementation.* Ann Acad Med Singapore, 2003. **32**(4): p. 466-76.
383. Tan, P.H., et al., *Screen detected breast cancer in an Asian population: pathological findings of the Singapore breast screening project.* Breast, 1999. **8**(3): p. 120-5.
384. Ng, E.H., et al., *Results of intermediate measures from a population-based, randomized trial of mammographic screening prevalence and detection of breast carcinoma among Asian women: the Singapore Breast Screening Project.* Cancer, 1998. **82**(8): p. 1521-8.
385. *SEER Stat Fact Sheets: Breast Cancer*. 20/11/2014]; Available from: http://seer.cancer.gov/statfacts/html/breast.html.
386. Guy, G.P., Jr., et al., *Economic burden of cancer survivorship among adults in the United States.* J Clin Oncol, 2013. **31**(30): p. 3749-57.
387. Basch, E., et al., *Recommendations for incorporating patient-reported outcomes into clinical comparative effectiveness research in adult oncology.* J Clin Oncol, 2012. **30**(34): p. 4249-55.
388. Gotay, C., *Core symptom measures in cancer clinical trials.* J Natl Cancer Inst, 2014. **106**(7).
389. Calvert, M., et al., *Patient-reported outcome (PRO) assessment in clinical trials: a systematic review of guidance for trial protocol writers.* PLoS One, 2014. **9**(10): p. e110216.
390. Snyder, C.F., *Using patient-reported outcomes in clinical practice: a promising approach?* J Clin Oncol, 2014. **32**(11): p. 1099-100.