uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Garth Nicholas

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Epidemiology)

GRADE / DEGREE

Department of Epidemiology and Community Medicine

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Survival Analysis of Patients Seen at the Ottawa Hospital Regional Cancer Centre with Early Breast Cancer 1985-2001: Effect of Changes in Stage and Adjuvant Chemotherapy over Time

TITRE DE LA THÈSE / TITLE OF THESIS

Carl van Walraven

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

Rama Nair

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

J. Little

B. Wilson

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Survival Analysis of Patients Seen at the Ottawa Hospital Regional Cancer Centre with Early Breast Cancer 1985-2001: Effect of Changes in Stage and Adjuvant Chemotherapy over Time

Garth Nicholas

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the degree of

Master of Science
in
Epidemiology

Department of Epidemiology and Community Medicine
Faculty of Medicine
University of Ottawa

Your file Votre référence
ISBN: 978-0-494-34096-7
Our file Notre référence
ISBN: 978-0-494-34096-7

Abstract

Breast cancer is the commonest cancer affecting Canadian women. Recent decades have seen a trend of decreasing breast cancer mortality. This has been attributed to several causes, including greater use of screening mammography and increased efficacy of adjuvant chemotherapy. We studied the effect of these two factors on the overall survival of breast cancer patients at the Ottawa Hospital Regional Cancer Centre. Data were collected from 2985 charts from the years 1985, 1988, 1992, 1995, 1998, and 2001. Adjuvant chemotherapy was associated with a decreased hazard of death from any cause (HR=0.783 p=0.0208). A decrease in mean tumour size seen over the time period of the study is potentially attributable to mammographic screening. Decreased tumour size was associated with a decreased hazard for death (HR 0.986 $p<0.0001$). No difference in overall survival between earlier and later cohorts could be demonstrated, perhaps due to shorter follow up in later cohorts.

Contents

List of Tables

List of Figures

1.0 Introduction

1.1 Trends of Breast Cancer Incidence and Mortality in Canada

Breast cancer is the most common cancer affecting Canadian women, with an estimated 22 300 new diagnoses and 5300 deaths in 2006.[1] The age-standardized incidence of breast cancer in Canada increased steadily from 82.2 per 100 000 women in 1973 to 102 per 100 000 in 1992 and has remained stable since then. However, breast cancer mortality has been decreasing. In 1990 the breast cancer mortality rate was 31.3 per 100 000. The current rate of 25.6 per 100 000 is the lowest since 1950.[1] This decrease in breast-cancer related mortality has also been seen in both the United States and the United Kingdom.[2]

The reason for this decline in mortality is not understood, and has been the source of some controversy.[2,3] Two important factors have been proposed to explain this decline. These are the widespread adoption of screening mammography for women above the age of 50, and improvements in adjuvant systemic therapy for women with early breast cancer. Use of screening mammography and adjuvant chemotherapy became common practice around the same time, so it has been difficult to determine the relative contribution of either to the observed declining mortality. The following sections briefly summarize the literature regarding the effect of screening mammography and systemic adjuvant therapy with both chemotherapy and hormonal therapy on patient outcomes.

1.2 The Effect of Screening Mammography on Patient Outcomes

Screening mammography is the provision of mammograms to asymptomatic women with the intent of discovering breast cancer before it becomes clinically evident.

Underlying screening mammography is the assumption that breast cancers discovered at an earlier stage in their natural history have a greater chance of being cured.

In Canada, screening mammography became commonplace in the mid 1980's, and continued to become more prevalent through the 1990's.[4,5,6] In 1988 Canadian guidelines for screening mammography recommended that women aged 50 to 69 years of age be offered mammography every two years as a component of a breast screening program that also includes breast exam by a health professional and teaching about breast self-exam.[7]

These recommendations have been widely adopted. Between 1988 and 2000, the number of women seen annually in Canadian organized breast screening programs increased from 4475 to 679 257.[7] This likely undercounts screening activity, since many women obtain their screening mammograms from their family doctor (i.e. outside the context of an organized screening program). Depending on the province, between 25% and 75% of screening mammograms are delivered through the fee-for-service sector by private physicians.[7]

Ontario guidelines for screening mammography are slightly different from the Canadian guideline in that they advocate screening mammography biannually up to the age of 74. Compliance with this recommendation is good, with the 2000-2001 Canadian Community Health Survey indicating that 65% of Ontario women in the target age range self-report having had a mammogram in the last two years.[8] In Ontario the proportion of screening mammograms performed outside of the provincial screening program was approximately 73% in 2001.[8]

Screening mammography has been studied extensively. Eight large randomized controlled trials compare the effect of screening mammography versus periodic clinical follow up on breast cancer-specific and overall mortality (Table 1.1).[9-16] With the exception of the Canadian studies (in which volunteers were enrolled at academic centres), all of these studies are community-based. Lists of a population were prepared and women were randomized by one means or another to be either invited or not invited to a mammography clinic. Mortality rates were determined using national mortality databases. Five of these studies show a trend towards decreased breast cancer-specific mortality, and two of these achieve statistical significance. The two Canadian studies are notable exceptions showing no effect of screening mammography on breast cancer mortality.

All of these studies have been criticized for their methodologies. In particular, the Scottish study has been critiqued for having too few clusters, and for an imbalance of

Table 1.1. Randomized trials of Screening Mammography

First Author Country	Publi-cation Year	Number Of Patients	Randomization Unit	Patient Ages	Frequency Of Screening	Control Intervention	Breast Cancer-Specific Mortality	All Cause Mortality
Chu KC, USA	1988	60696	Matched Pairs	40-64	Annual	Usual Practice	RR at 14 years 0.78 (0.63-0.96)	NR
Andersson I, Sweden	1988	42283	Individual	44-69	18-24 months	Usual practice	RR at 8 years 0.96 (0.68-1.35)	RR at 8 years 0.99 (0.93-1.05)
Tabar L, Sweden	1989	133092	Cluster	40-74	Biannual	Usual practice	RR at 10 years 0.69 (0.55-0.88)	RR at 10 years 0.99 (0.95-1.03)
Frisell J, Sweden	1991	40318	Individual*	40-64	Biannual	Usual practice	RR at 11 years 0.74 (0.5-1.1)	NR
Miller AB, Canada	1992	50430	Individual	40-49	Annual***	Annual clinical exam	RR at 16 years 1.06 (0.8-1.4)	NR
Miller AB, Canada	1992	39405	Individual	50-59	Annual***	Annual clinical exam	RR at 13 years 1.03 (0.78-1.33)	NR
Alexander F Scotland	1994	44288	Cluster	45-64	Biannual	Usual practice	RR at 14 years 0.79 (0.6-1.02)	RR at 10 years 0.85 (0.79-0.92)
Bjurstam N, Sweden	2003	51611	Individual**	39-69	18 months	Single screen	RR at 14 years 0.79 (0.58-1.08)	NR

*Individuals were "randomized" by their date of birth, with women born between the 11th and 20th of the month enrolled to the control arm, and all others enrolled on the screening arm

**18% of patients were "randomized" by birth date as described above

***The screening strategy in the Canadian studies consisted of annual mammography and clinical breast exam

baseline prognostic variables in these clusters. More of the clusters randomized to mammography were in areas of high socioeconomic status, perhaps accounting for the surprising finding that the mammography group had no improvement in breast cancer-specific survival, but had a significant improvement in overall survival when compared to the non-mammography group. These problems have been extensively reported by the study author[21] and this trial is seldom considered when the evidence for mammography is reviewed.[22] In fact, it is considered bad luck to mention this study by name, thus its invariable sobriquet, 'the Scottish study'.

The American Health Insurance Plan (HIP) study[12] used a matched-pair randomization where pairs of women were matched on certain baseline factors prior to one member of the pair being randomized to mammography. This method has been criticized because it created demonstrably unbalanced groups which one author claims are findings "incompatible with adequate randomization".[20] Furthermore, the study excludes women with previous breast cancer; however there is evidence that information about previous diagnoses was collected much more thoroughly on screened patients than on control patients, potentially biasing the study in favour of the screened group.

The two trials which "randomized" patients based on their date of birth [13,14] did not have adequate concealment of randomization, thus introducing potential bias. This bias would be a fatal flaw for most clinical trials enrolling volunteers, but may be less important in the context of a trial enrolling the entire population of a particular region.

The Canadian studies have been heavily criticized, despite being the most methodologically sound. Some have claimed that the annual clinical exam mandated for the control group is a cointervention which dilutes the effect of mammography. Indeed,

the breast cancer diagnosis rate in the control groups of these Canadian trials is similar to the rate found in some European countries that routinely use screening mammography .[22] The Canadian trials have also been criticized for enrolling volunteers who presented to academic health care centres, rather than the more community-based enrollment schemes of the European trials. Although this criticism is not groundless, the fact that the Canadian studies were adequately randomized means that an effect of mammography should be demonstrable, even in a study population which may not represent the general population.

Finally, all of the trials may be criticized for their use of breast cancer-specific mortality as the primary outcome. This outcome was selected to avoid both length-time and lead-time bias, both biases which tend to favour screening.[24] However, as compared to all-cause mortality, disease-specific mortality can be attributed incorrectly due to administrative or other problems with death certificates[23] or can be biased by knowledge of screening status. All-cause mortality would capture the deaths of patients who die as a consequence of the screening intervention (unlikely in this instance) and also places the mortality benefit of screening in the context of competing causes of mortality.
Although all of the trials used breast cancer-specific mortality as their primary end point, most of the studies also reported other end points that are relevant to this thesis. These include the size or stage of tumours detected by screening, and the survival of screened and non-screened patients after a diagnosis of breast cancer. These results are summarized in Table 1.2.

Table 1.2. Relevant Secondary Outcomes from Mammography Trials			
First Author And Country	**Tumour Size or Stage**		**Risk of mortality among patients diagnosed with breast cancer***
	Screened	**Unscreened**	
Chu KC, USA	56% Stage I	48% Stage I	RR at 14 years 0.79 (0.65-0.97)
Andersson I, Sweden	61% Stage I	41% Stage I	NR
Frisell J, Sweden	60% Stage I	52% Stage I	NR
Miller AB, Canada	51% <20 mm 79% node-negative	43%<20 mm 59% node-negative	NR
Miller AB, Canada	50% <20 mm 67% node-negative	34% <20 mm 58% node-negative	NR
Alexander F, Scotland	55% <20 mm	34% <20 mm	NR
Bjurstam N, Sweden	RR of node-positive disease in screened group 0.80 (0.61-1.05)		RR at 14 Years 0.83 (0.63-1.09)**

* Relative risks are risk of death in the screened group compared to the control group.
** Not reported, but calculated from data presented in the paper.

The risk of mortality among patients who are diagnosed with breast cancer is obviously suspect due to lead-time bias and length-time bias. It is for this reason that it is not reported routinely in most trials. There may be some true effect, though, because of the known prognostic significance of tumour size.[55] As Table 1.2 shows, the mammographically screened population has smaller tumours at diagnosis. This observation has also been made in a large Canadian community-based sample.[26]

Three published meta-analyses of these trials have concluded that screening mammography provides a survival benefit to women over the age of 50[17-19], while a fourth meta-analysis does not reach this conclusion.[20] The author of the dissenting meta-analysis also wrote the review for The Cochrane Collaboration and concluded that screening mammography does not improve all cause mortality in the general population, though breast-cancer specific mortality may be improved.[20] The meta-analyses are summarized in Table 1.3.

Table 1.3. Published meta-analyses of mammography trials				
Author	**Year**	**Studies Included**	**Patients Included**	**RR of breast cancer -specific death**
Nystrom L	1993	5 (All Swedish studies)	282777	At 12 years 0.77 (0.67-0.88)
Kerlikowske K	1995	8 RCT 4 Case-control	NR	At 7- 9 years 0.79 (0.66-0.83)
Gøtzsche PC	2000	3 (Canadian and one Swedish study)	132118	1.04 (0.84-1.27)
Hackshaw A	2003	6 (Canadian studies excluded)	372288	0.76 (0.67-0.87)

The different meta-analyses come to their different conclusions largely due to their decisions to include or exclude specific studies. The Nystrom meta-analysis was not intended to encompass all of the trials in the literature, but rather to summarize the five Swedish trials. The Kerlikowske study included data from some non-randomized trials, and also included every published RCT, including the widely disregarded Scottish study. The paper gives little evidence that studies were evaluated for their quality before being included. The author states that a test for heterogeneity was employed, but never discusses the results.

The Hackshaw paper represents the position of the European Society of Mastology (EUSOMA). It includes all randomized trials except for the Canadian trials. The author states that because the Canadian trials mandated clinical breast exam for patients in the control arm the studies were not designed to compare mammography to an unscreened group.

The most controversial meta-analysis is that by Gøtzsche. This paper excludes every study except for the Canadian studies and one Swedish study on the grounds that the other studies are methodologically flawed. The author takes specific issue with the methods and adequacy of randomization in most of the studies and demonstrates that

these deficiencies in randomization lead to experimental groups with significant imbalances at baseline. Many of these imbalances favour the screened groups. This review has prompted much debate and controversy but has not changed practice. If nothing else it emphasizes the need for sound trial design, particularly in a field as fraught with potential biases as population screening.

The issue of screening women less than 50 years of age has been reviewed in six different meta-analyses.[15,25-30] The conclusions of these reviews are divided evenly, with three of them concluding that there is no benefit in this population while the other three claim to demonstrate significant breast cancer-specific mortality reductions. There are relatively few women in this age group enrolled in the RCTs, and these small numbers probably account for the variance in conclusions reached by the different meta-analyses.

In summary, screening mammography is a common procedure which has an established place in the primary care of Canadian women. Despite this wide acceptance, there are questions about the effectiveness of this intervention. Most studies suggest an improvement in breast cancer-specific survival, though no trial or meta-analysis has compellingly suggested a benefit in all-cause mortality. Most studies demonstrate a tendency for screen-detected cancers to be smaller and earlier stage than those not detected by screening. Because of this mammography may convey a survival advantage in the subset of women who have been newly diagnosed with breast cancer.

1.3 The Effect of Adjuvant Systemic Therapy on Patient Outcomes

Most women with breast cancer present with disease that is clinically limited to the breast and the ipsilateral axillary lymph nodes (i.e. lymph nodes on the same side as

the tumour). This is termed early breast cancer (EBC). However, many women have their breast cancer recur elsewhere in their body after appropriate local treatment. This observation suggests that microscopic or occult metastatic disease was present in these women at the time of diagnosis. Adjuvant systemic therapy is administered to patients to eradicate micrometastatic disease and prevent systemic recurrence of cancer. Adjuvant systemic therapy consists of hormonal therapy and chemotherapy. These are administered either alone or in combination and are briefly summarized below.

1.3.1 Adjuvant Hormonal Therapy

Hormonal therapy was the first type of systemic therapy used to treat breast cancer. More than a century ago it was observed that patients with metastatic breast cancer occasionally had their disease go into remission after surgical oophorectomy.[31] Through the course of normal life and development human breast tissue goes through hormone-dependant changes related to puberty, pregnancy, lactation, and menopause. The cells of mammary glands have receptors for estrogen and progesterone, two hormones that work together to direct mammary epithelial growth, differentiation, and survival.[32] Some tumours arising from breast tissue retain this hormone sensitivity, and various methods of hormone manipulation have been studied for their ability to inhibit the formation or progression of these cancers. Specifically, therapies attempt to deprive tumour cells of estrogen, which acts as a growth factor for hormone-sensitive tumours. Three broad categories of hormonal manipulation include ovarian suppression, aromatase inhibition, and the use of selective estrogen receptor modulators (SERMs) such as tamoxifen.

1.3.1.1 Ovarian suppression

The ovaries are the major source of estrogen in pre-menopausal women. Ovarian suppression or ablation has been achieved by surgery (oophorectomy), medicine (treatment with GnRH agonists), and radiation.

There are 13 randomized trials of either surgical or radiation ablation. Five of these trials included chemotherapy for both arms, with ovarian ablation added to the experimental group. The trials are summarized in Table 1.4.

The interesting finding in this table, other than the reduction in risk of death associated with ovarian ablation, is the fact that the trials comparing ovarian ablation with chemotherapy to ovarian ablation alone show similar survival advantage. A potential explanation for this is a common mechanism of action for chemotherapy and

Table 1. 4: Trials of surgical or radiation ablation of ovaries as adjuvant therapy for breast cancer

Trial	Year	Ovarian Treatment	Systemic Treatment	Number Enrolled*	RR of Death At 15 years
Christie	1959	Radiation	None	178	0.97(0.87-1.08)
Norwegian	1991	Radiation	None	151	0.68(0.46-0.99)
NSABP B-03	1970	Surgery	None	184	0.91(0.71-1.17)
Saskatchewan	1981	Surgery	None	255	0.75(0.58-0.96)
PMH Toronto	1996	Radiation	Prednisone	349	0.88(0.76-1.00)
Ontario CTRF	1977	Radiation	None	9	0.83(0.58-1.19)
CRFB Caen A	1976	Radiation	None	1	
Rhode Island Stratum 1	1985	Surgery	None	42	0.50(0.24-1.72)
Rhode Island Stratum 2	1985	Surgery	MTt	38	1.01(0.32-3.19)
Toronto-Edmonton	1988	Radiation	CMF+BCG	241	1.10(0.83-1.46)
Vancouver	1988	Radiation	CMF	111	0.95(0.59-1.52)
IBCSG	1985	Surgery	CMFPr	281	0.95(0.80-1.14)
SWOG	1989	Surgery	CMFVPr	262	0.79(0.57-1.11)
Total				**2102**	**0.89(0.82-0.96)**
Total, non-chemo trials				**1169**	**0.85(0.78-0.93)**
Total, chemo + OA trials				**933**	**0.94(0.81-1.08)**

Adapted from reference 36.

*number of patients <50 years of age. Because few of the trials required objective evidence that patients were menopausal, and because ovarian ablation is not expected to be efficacious in menopausal women, women over 50 years of age are excluded.

ovarian ablation. It may be that much of the effect of cytotoxic chemotherapy comes not from the chemotherapy *per se*, but from the chemical ovarian ablation which often results from chemotherapy administration.

Ovarian ablation can also be achieved medically, through the use of a class of medications called gonadotropin-releasing hormone (GnRH) agonists. These medications inhibit the release of GnRH from the pituitary gland. GnRH normally controls the activity of the ovaries and its inhibition results in a sharp decrease in levels of circulating estrogen in premenopausal women.

Nine randomized trials have studied the adjuvant use of GnRH agonists. Five of them compare the hormonal therapy to chemotherapy, while the other three included chemotherapy for both arms and GnRH agonist added to the experimental group. These trials are summarized in Table 1.5.

Table 1.5: Trials of GnRH agonists in the adjuvant setting

First Author	Year	Hormone Therapy	Chemotherapy	N	Outcome
Trials comparing GnRH agonists to chemotherapy					
Kaufmann[33]	2001	Goserelin	CMF	1640	HR of death at 6 years 1.01 (0.84-1.20)*
Boccardo[34]	2000	Goserelin Tamoxifen	CMF	244	HR of death at 7 years 0.71 (0.38-1.31)
Jakesz[35]	2001	Goserelin Tamoxifen	CMF	1189	HR death at 5 years 0.99 (0.76-1.28)*
Roche[36]	2000	Triopterelin Tamoxifen	FEC	333	Overall survival 4.5 years No difference p=0.18**
Wallweiner[37]	2002	Leoprorelin	CMF	227	At 31 months, no difference In progression-free survival (p=0.94)
Trials that Added GnRH agonists to Chemotherapy					
Baum[38]	2002	Goserelin	Usual care***	2648	HR of death at 5.5 years 0.78 (0.60-1.03)**
Davidson[39]	1999	Goserelin ±Tamoxifen	CAF	1504	No benefit with addition of goserelin to chemo
IBCSG[41]	2003	Goserelin	CMF	1063	RR of recurrence at 8 years 1.13 (0.83-1.53)
DeMatteis[40]	1998	Goserelin	Epirubicin	92	No additional benefit with Addition of hormones**

*ER positive patients only
** presented in abstract only
*** whatever combination of surgery, chemotherapy, and tamoxifen the patient's oncologist thought to be appropriate

Overall, these trials indicate an approximate equivalence between ovarian suppression and chemotherapy (usually CMF) in the adjuvant setting for women with hormone-responsive cancers. The addition of GnRH agonists to chemotherapy seems to add little. This is a similar finding to the trials which added surgical or radiation ablation to chemotherapy.

Despite the presence of such a large number of randomized trials, ovarian ablation is much less commonly used than chemotherapy in North America.[42] Many oncologists point out that CMF chemotherapy is no longer the standard of care, and that there are few trials comparing ovarian suppression to more modern regimens containing anthracyclines. As the results of more trials become available in the coming years, ovarian ablation may take on more prominence.

1.3.1.2 Aromatase inhibition

In the post-menopausal woman the ovaries no longer produce significant quantities of estrogen. In this circumstance, circulating estrogen is derived from androgens generated by the adrenal glands. Aromatase is an enzyme found in fat, muscle and adrenal glands that catalyzes the conversion of androgens to estrogens. This is the major source of estrogen in the post-menopausal female and provides the theoretical justification for aromatase inhibition as a treatment for breast cancer in this population.

Aromatase inhibitors that are clinically available include exemestane, anastrozole, and letrozole. Although there has long been data supporting the use of these agents in patients with metastatic disease[46-49], by 2001 there were only a handful of trials looking at these agents in the adjuvant setting, either alone or in sequence or combination with

tamoxifen[43-45]. Because the data on aromatase inhibition as an adjuvant therapy is quite recent, very few of the patients included in this thesis will have received adjuvant aromatase inhibitors. The data supporting their use will not be presented in any more detail.

1.3.1.3 Selective Estrogen Receptor Modulators (SERMs)

SERMs are drugs which act as analogs of estrogen. They occupy the estrogen-binding site of the estrogen receptor, but do not stimulate the downstream mechanisms that promote cell division. Most of the drugs in this class have some weak estrogenic effects in certain tissues in addition to their predominant anti-estrogen properties. This characteristic explains why these drugs are used to treat osteoporosis, and why they are associated with the development of endometrial cancers.[51] Several SERMs are available, including tamoxifen (Nolvadex), raloxifene, and torimifene. The only SERM relevant to the adjuvant treatment of breast cancer during the time covered by this thesis is tamoxifen.

There have been more than 70 randomized trials of adjuvant tamoxifen versus either placebo or some other intervention. These will not all be reviewed in detail. The pivotal trials will be highlighted, and then the entire literature summarized in a meta-analysis.

The first trial to show a survival advantage associated with adjuvant tamoxifen was the trial conducted by the Nolvadex Adjuvant Treatment Organization (NATO trial), published in 1985.[52] This British study enrolled 1131 women, both pre-and post-menopausal, and randomized them to two years of tamoxifen therapy after surgery or to surgery alone. The initial report, after 6 years of follow-up, showed a relative reduction

in overall mortality of 30%. An advantage in recurrence-free survival was also shown. By 1995, with a mean of 16 years of follow-up, there had been 249 deaths in the 563 women in the tamoxifen arm and 292 deaths in the 568 women who did not receive tamoxifen (OR for death 0.78 (0.62-0.98)).[53]

After the NATO trial, several smaller trials followed which replicated NATO's finding of improved progression-free survival, but which were probably underpowered to show a difference in overall survival. Confirmation of the survival advantage conveyed by tamoxifen came with a 1987 publication from the Scottish National Trial Group.[54] This study randomized 1312 women to either five years of tamoxifen or no tamoxifen after surgery. After eight years of follow-up, the overall survival favoured the tamoxifen group (HR for death 0.71 [0.58-0.89]).

At the time, there was considerable controversy regarding which patient subsets might derive most benefit from tamoxifen. Specifically, there was interest in whether the benefit might be greater in pre- or post-menopausal women, and in women whose tumours did or did not involve local lymph nodes ("node-positive" or "node-negative"). The Scottish trial suggested that the benefit was greatest in women with node-positive disease, and in post-menopausal women. In fact, the hazard ratio for death was not statistically significant for pre-menopausal women (HR 0.57 [0.27-1.19]). The NATO trial did not enroll enough pre-menopausal women for a credible subgroup analysis. The first meta-analysis of these trials suggested that the benefit of tamoxifen was limited to post-menopausal women.[55]

Table 1.6. Results of EBCTG Overview of adjuvant tamoxifen			
Group	**Years of Tamoxifen**	**Outcome**	**Peto Odds Ratio (99% CI)**
All Patients	1-5	Overall survival at 10 years	0.85 (0.82-0.89)
All Patients	1-5	Breast Cancer Recurrence	0.74 (0.71-0.76)

Subsequent trials benefited from improved ability to detect whether a given tumour did or did not express estrogen receptors ("ER-positive" or "ER-negative"). Tumours which are ER-negative constitute about one-third of breast cancers, and would not be expected to be as sensitive to hormonal manipulation as ER-positive tumours. Therefore, inclusion of women with ER-negative tumours in trials might lead to systematic underestimation of the benefit of tamoxifen. This might particularly be the case in pre-menopausal women, who have a higher incidence of ER-negative tumours than women who are post-menopausal.

The Early Breast Cancer Trialists Collaborative Group (EBCTCG) performs a meta-analysis with individual patient data of randomized trials of tamoxifen every five years. The first was published in 1990,[55] and the most recent update in 2003.[53] The 2003 update includes 37099 women in 55 trials, all started before 1990. All trials compare adjuvant tamoxifen to no adjuvant hormonal therapy. The primary outcomes are summarized in Table 1.6, while significant subgroup analyses are found in Table 1.7.

The subgroup analyses identify several interesting points. Firstly, five years of tamoxifen seems superior to shorter courses in all subgroups. Secondly, there is very little benefit to tamoxifen for women with ER-negative tumours. The small observed advantage may exist because ER-negative tumours seldom have no estrogen receptors whatsoever, but rather express them at a level below an arbitrarily defined cut off.

Table 1.7 EBCTG Overview of adjuvant tamoxifen: subgroup analyses

Group*	Years of Tamoxifen	Overall survival at 10 years		Breast Cancer Recurrence	
		Relative Risk (95% CI)	Absolute Risk Reduction	Relative Risk (95% CI)	Absolute Risk Reduction
ER-positive	1	0.86 (0.852-0.867)	5.2%	0.79 (0.782-0.798)	6.3%
	2	0.82 (0.813-0.827)	4.6%	0.72 (0.714-0.726)	7.6%
	5	0.72 (0.710-0.730)	5.9%	0.50 (0.488-0.512)	14.9%
ER-negative	1	0.94 (0.938-0.941)	2%	0.94 (0.936-0.944)	1.1%
	2	0.93 (0.927-0.933)	0.5%	0.87 (0.865-0.875)	2.3%
	5	0.97 (0.967-0.973)	-3.4%	0.94 (0.935-0.945)	1.3%
Node-Negative	1	0.87 (0.862-0.878)	3.4%	0.83 (0.820-0.841)	3.4%
	2	0.89 (0.885-0.895)	2.3%	0.72 (0.710-0.730)	5.7%
	5	0.75 (0.741-0.759)	5.6%	0.51 (0.498-0.522)	14.9%
Node-positive	1	0.88 (0.876-0.884)	4.5%	0.79 (0.785-0.795)	7.7%
	2	0.81 (0.806-0.814)	7.2%	0.70 (0.694-0.706)	9.2%
	5	0.72 (0.708-0.732)	10.9%	0.57 (0.557-0.583)	15.2%
Age <50	5	0.68 (0.658-0.703)	6.7%	0.55 (0.528-0.573)	14.1%
Age 50-59	5	0.89 (0.883-0.897)	1.4%	0.63 (0.615-0.645)	9.8%
Age 60-69	5	0.67 (0.656-0.684)	7.6%	0.46 (0.445-0.475)	16.1%
Age >70	5	0.66 (0.630-0.691)	9.7%	0.46 (0.422-0.501)	16.8%

* Except for the group identified as "ER-negative", all groups exclude known ER-negative patients
Some data adapted from reference 56.

Third, for women with ER-positive tumours the relative mortality benefit of five years of tamoxifen is roughly 30% in all age groups, supporting the theory that early trials were negative in younger women because this population includes a larger number of ER-negative tumours. Finally, the overview shows that the relative benefit of tamoxifen is the same in both node-negative and node-positive patients, though the absolute benefit is greater in those who are node-positive. This is because node-positive women have a higher baseline risk for recurrence and breast cancer-related death.

The indications for tamoxifen have changed over time as our understanding of this medication has evolved. In the late 1980's tamoxifen was first recommended for post-menopausal women who were node-positive. Current Ontario[57] and Canadian[58,59] practice guidelines recommend that tamoxifen should be a component of the adjuvant treatment of virtually all women with ER-positive tumours, whether pre- or post-menopausal.

1.3.2 Adjuvant Chemotherapy

Chemotherapy is a term describing a group of more than fifty drugs used to kill cancer cells. Most of these drugs act by disrupting a cell's ability to replicate build or replicate DNA. DNA replication is a critical component of cell division. It was once thought that cancer cells grow and divide more rapidly than normal cells, and that cancer cells should therefore be differentially susceptible to drugs which disrupt the machinery required for cell growth. Although this assumption about the rate of cancer cell division has proven to be true in only a minority of tumour types, many chemotherapy agents have been found which are clinically useful.

There are more than 100 reported randomized trials of adjuvant chemotherapy for breast cancer.[60] Not all will be discussed. Instead, we will focus on the pivotal studies which established the efficacy of chemotherapy in the adjuvant treatment of breast cancer, or which led to the adoption of the anthracycline-based regimens which form the standard of care today. Some preliminary data about adjuvant use of taxanes was available by 2001, and will be briefly reviewed. Finally, an EBCTCG meta-analysis of relevant trials will be summarized at the end.[59] Developments which came after 2001, such as dose-dense adjuvant chemotherapy, sequential hormonal therapies, or the addition of trastuzumab to chemotherapy are not be relevant to the patients in this thesis and will not be reviewed.

1.3.2.1 Early Trials and CMF

The first randomized trials of prolonged adjuvant chemotherapy after surgery for breast cancer were started in 1972 and 1973 in the United States and Italy, respectively.[62,63] The American study used oral l-phenylalanine mustard (L-PAM) and showed statistically significant improvements in the proportion of patients disease-free and alive at 10 years. The L-PAM adjuvant regimen has never been widely used in practice, however, because its outcomes were not as good as those seen in the Italian trial.

The Italian study used a combination of three chemotherapy drugs: cyclophosphamide, methotrexate, and 5-fluorouracil (CMF). The latter two drugs were given intravenously on days 1 and 8 of each month, while the cyclophosphamide was given orally on days 1 through 14. This treatment was repeated for 12 months. This regimen came to be known as "Classical CMF" or "Bonadonna CMF" (after the first author of the paper) to distinguish it from later CMF variants which administered the cyclophosphamide intravenously.

The Italian trial enrolled only women with node-positive breast cancer. Women were either pre- or post-menopausal. Key results are summarized in Table 1.8. This trial established several points which were to become axiomatic for years in the adjuvant treatment of women with breast cancer. The most important observation was that chemotherapy seemed to be more effective in pre-menopausal women. There are many theories as to why this may be, and the most compelling relates to the effect of chemotherapy on the ovaries of pre-menopausal women, as mentioned previously.

Table 1.8. Results of Italian CMF study

	Disease Free at 10 years (%)			Alive at 10 years (%)		
Group	**Control**	**CMF**	**P**	**Control**	**CMF**	**p**
All patients	31	43	0.001	47	55	0.10
Post-menopausal	32	38	0.32	50	52	0.89
Pre-menopausal	31	48	0.0005	45	59	0.02
Pre-menopausal 1-3 nodes	40	61	0.0002	51	68	0.025
Pre-menopausal >3 nodes	15	26	0.03	30	42	0.29

Following the positive results of the Italian trial, a second generation of studies was published which confirmed the important findings regarding overall- and disease-free survival. These trials also looked at the appropriate duration of therapy and found that six cycles of CMF was equivalent to twelve.[64,65]

1.3.2.2 Anthracycline-based Regimens

The next advance in adjuvant therapy was the addition of anthracyclines to adjuvant chemotherapy regimens. The anthracyclines are a family of drugs including doxorubicin (Adriamycin), epirubicin, daunorubicin, and idarubicin. Only the first two are used in the treatment of breast cancer. They work mainly by inhibiting the activity of topoisomerase II, an enzyme required for DNA replication.[66] In the 1980's, anthracycline-based regimens were found to be the most active chemotherapy regimens against metastatic breast cancer.[66,67] This led investigators to consider the addition of these drugs to adjuvant regimens.

The first significant study was a French study which looked at a combination called AVCF (doxorubicin, vincristine, cyclophosphamide, 5-fluorouracil) versus CMF in node-positive patients. The initially reported results[68] did not demonstrate a survival advantage for the anthracycline-containing regimen, though later analysis of more mature

Table 1.9: Most commonly used anthracycline regimens			
Regimen	FEC [72]	FAC [71]	CEF [73]
Fluorouracil dose	600 mg/m^2 day 1 & 8	500 mg/m^2 day 1	500 mg/m^2 day 1 & 8
Cyclophosphamide dose	600 mg/m^2 day 1 & 8	500 mg/m^2 day 1	75 mg/m^2 day 1-14
Anthracycline and dose	Epirubicin 50 mg/m^2 day 1	Doxorubicin 50 mg/m^2 day 1	Epirubicin 60 mg/m^2 day 1& 8
Survival advantage versus CMF	87% vs. 74% at 5 years (p=0.06)	75% vs. 69% at 5 years (p=0.18)	77% vs. 70% at 5 years (p=0.03)

data (at 16 years) showed that 56% of the AVCF arm were still alive, compared to 41% of women in the CMF arm (p=0.01).[69] This trial, along with an American study,[70] led the way to three studies which established the anthracycline-based regimens used today.

The three most commonly used anthracycline-based regimens[71-73] are summarized in Table 1.9. They all include 5-fluorouracil, cyclophosphamide, and either doxorubicin or epirubicin. The regimens vary slightly by drug doses and schedules. Each of these regimens has been found to be superior to CMF in terms of disease-free survival, though the results for overall survival have not been uniformly positive in single clinical trials.

The EBCTCG has done a meta-analysis of 69 trials started before 1990 which look at the issue of adjuvant chemotherapy.[61] Like the tamoxifen meta-analysis,[53] this is updated every five years. The most recent analysis is based on individual patient data for 18718 women. It includes 47 trials with a "chemotherapy versus no chemotherapy" design, 11 trials looking at longer versus shorter courses of chemotherapy, and 11 trials looking at regimens containing anthracyclines compared to CMF. The primary outcomes of this meta-analysis are summarized in Table 1.10, while important subgroups are addressed in Table 1.11 and Figure 1.1.

These results indicate that the relative risk reduction is similar across all groups, regardless of receptor status, nodal status, or age, though there is a trend towards decreasing benefit with increasing age. Of course, the absolute risk reduction varies

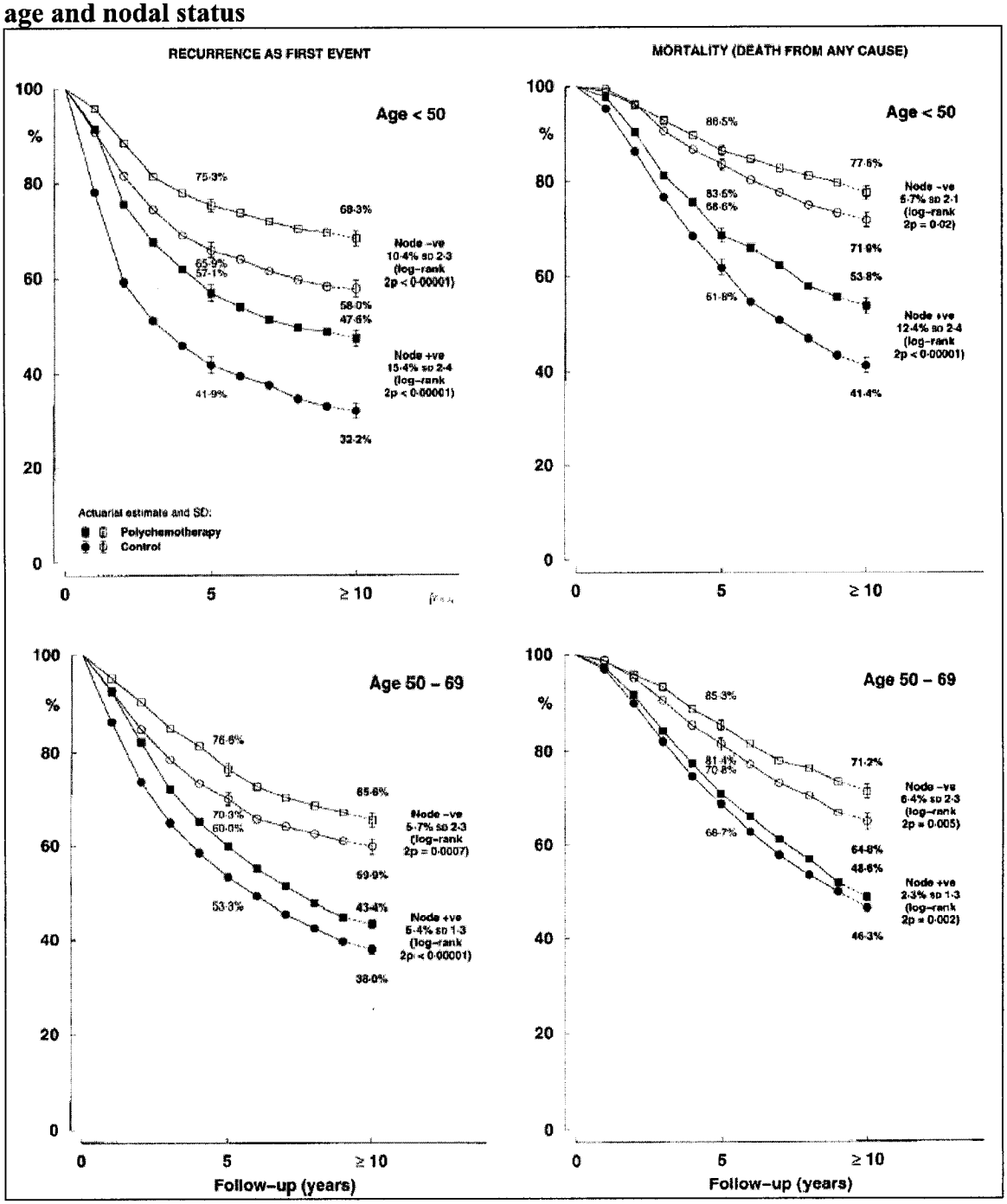
Table 1.10. Primary outcomes of the EBCTCG overview of adjuvant chemotherapy			
Comparison	N	Odds Ratio of Recurrence* (99% CI)	Odds Ratio of Death* (99% CI)
Chemo vs. no chemo	17723	0.76 (0.73-0.8)	0.85 (0.80-0.89)
Anthracycline vs. CMF	5942	0.88 (0.81-0.97)	0.89 (0.80-0.98)

*At 10 years

Table 1.11. EBCTCG overview of adjuvant chemotherapy: subgroup analyses				
	10-year Overall Survival		**Breast Cancer Recurrence**	
Group*	**Relative Risk (95% CI)**	**Absolute Risk Reduction**	**Relative Risk (95% CI)**	**Absolute Risk Reduction**
Age <40	0.73 (0.715-0.746)	5.4%	0.63(0.613-0.648)	9.1%
Age 40-50	0.73 (0.720-0.740)	7.7%	0.66 (0.648-0.672)	10.0%
Age 50-60	0.86 (0.856-0.864)	3.3%	0.78 (0.773-0.787)	5.8%
Age 60-70	0.92 (0.917-0.923)	2.3%	0.82 (0.814-0.826)	5.5%
Age<50 ER Positive	0.80 (0.785-0.815)	6.6%	0.67 (0.652-0.689)	11.2%
Age <50 ER Negative	0.65 (0.638-0.662)	8.8%	0.60 (0.582-0.619)	13.2%
Age 50-70 ER Positive	0.81 (0.803-0.817)	2.3%	0.82 (0.814-0.826)	4.8%
Age 50-70 ER Negative	0.83 (0.822-0.838)	4.6%	0.70 (0.689-0.711)	8.3%

Some data from reference 74.

Figure 1.1. Outcomes of the EBCTG overview of polychemotherapy, stratified by age and nodal status



From reference 74.

based on the baseline level of risk. This suggests that the early trials showing benefit restricted to node-positive pre-menopausal women (the highest risk group) may have been underpowered to show benefit in other subgroups.

Current Canadian[58,59] guidelines suggest that all node-positive and selected node-negative women should be offered adjuvant chemotherapy, with or without tamoxifen. The guidelines acknowledge that there is little evidence for chemotherapy in women older than 70 years of age, due to small numbers of such patients enrolled in trials.

1.3.2.3 Beyond Anthracyclines

No sooner had anthracycline-containing regimens become the standard of care than investigators began looking to add other active drugs to adjuvant regimens. Most work has been done with a class of drugs called the taxanes, which include the agents docetaxel (taxotere) and paclitaxel (taxol). Both of these are known to be active in metastatic breast cancer.[75] Their mechanism of action involves interference with the activity of microtubules: intracellular organelles which are essential for cells undergoing mitosis.

The optimal integration of taxanes into adjuvant therapy is the subject of ongoing research and several trials which had not yet reported results by 2001, the last year covered in this thesis. Although there are now multiple mature phase III trials examining the integration of taxanes into an adjuvant regimen,[76-78] only one[76] was available early enough to have a potential impact on the patients studied for this thesis. This trial compared four cycles of adriamycin-cyclophosphamide (AC) to a similar regimen followed by four cycles of paclitaxel (AC-T). Although this trial did not publish its final

results until 2003, preliminary results suggesting a survival advantage for the AC-T arm were presented at meetings as early as 1999.[79] This trial enrolled 3121 women with node-positive breast cancer. At seven years of follow up, 74% of women in the AC-T arm were alive, compared to 68% in the AC arm (p=0.0064). Comparing the paclitaxel to the no-paclitaxel group, the hazard ratio for death was 0.82 (0.71-0.95).

1.4 Population-Based Survival Analyses in Breast Cancer

Survival analyses based on several cancer registries all over the world have been published. (Table 1.12) Of the North American Studies, most are from the United States and use the Surveillance, Epidemiology, and End Results Database (SEER) of the National Cancer Institute. This database was instituted in 1973 and, at that time, included cancer registries from nine American states. It has since grown to include more states and since 2000 has covered 26% of the American population. Although it is a very important resource, SEER has its limitations. SEER has recorded some stage information since its inception, but only began to collect information about tumour size and lymph node status in 1988. It did not collect information about steroid-hormone receptor status until 1990.[81] Although SEER contains good information about surgery[82] and radiation therapy[83] systemic therapy is "substantially underreported"[84] and not available in SEER public use files. The fact that SEER contains incomplete information about chemotherapy leads to significant gaps in the information which can be obtained. It is difficult to assess the impact of systemic therapy in general, and impossible to measure the effect of individual therapeutic approaches as they have evolved over time.

There are a large number of recently published reports based on the SEER database (see Table 1.12) correlating with SEER's recent decision to make many of their

files public and to provide custom software to assist with data manipulation. Of the published survival analyses using SEER data, none explicitly examine changes in stage

Table 1.12: Characteristics of Published Population-Based Survival Analyses of Patients with Breast Cancer

First Author	Country	Year	Years Studied	Stages Studied	Ages Studied	Sample Size	Stage Data	Screening Data	Chemo Data
Ugnat[99]	Canada	2004	1994-1997	All	All	2192	Yes[g]	No	Yes[h]
Pisani[103]	Great Britain	2004	1983-1998	All	All	20489	For 52%	No	Yes[h]
Chia[104]	Singapore	2004	1968-1992	All	All	6397	Yes	No	No
Jensen[105]	Denmark, Sweden	2003	1996-1997	All	All	2104	Yes	Yes	No
Chia[96]	Canada	2003	1991-2001	IV	All	2152	Yes[c]	No	No
Gaffney[81]	USA	2003	1988-1997	All	All	45880	Yes[a]	No	No
Henson[95]	USA	2003	1992-1999	All	All	79690	Yes[a]	No	No
Maggard[85]	USA	2003	1992-1998	All	<35 and 50-55	24935	Yes[a]	No	No
Brenner[106]	Finland	2003	1958-	All	<50	-[f]	Yes[b]	No	No
Li[89]	USA	2003	1992-1998	All	All	124934	Yes[a]	No	No
Li[93]	USA	2003	1974-1998	All	All	164958	No	No	No
Shavers[87]	USA	2003	1990-1997	All	<35	3978	Yes[a]	No	No
Clegg[90]	USA	2002	1975-1997	All	All	1780000	Yes[a]	No	No
Chaudhry[94]	Canada	2001	1991	I	All	938	Yes[c]	No	No
Diab[86]	USA	2000	1973-2000	All	All	307115	Yes[a]	No	Yes[d]
Paszat[85]	Canada	2000	1982-1991	All	All	46687	No	No	No
McCarthy[88]	USA	2000	1987-1993	All	>67	4438	Yes[a]	Yes	No
Olivotto[25]	Canada	1999	1989-1996	All	40-90	10095	Yes	Yes	Yes[e]
Chang[92]	USA	1998	1975-1992	IIIB	All	181083	Yes[c]	No	No
Henson[83]	USA	1995	1983-1987	All	All	56268	Yes[a]	No	No
Olivotto[88]	Canada	1994	1974-1984	All	All	3263	No	No	Yes[e]
Proposed	Canada		1985-2001	I-IIIB	All	Approx 3500	Yes	No	Yes

a. SEER database contains staging information which is slightly modified from the standard TNM format. Although it started in 1974, SEER did not record tumour size or lymph node status until 1988, and did not record steroid-hormone receptor status prior to 1990.
b. Staged cancers as localized or regional, rather than using formal staging.
c. All patients in the study were a particular stage.
d. See introduction for a discussion of the limitations of chemotherapy data in SEER.
e. Chemotherapy information was not collected at the level of individual patients.
f. Abstract only. Sample Size not stated.
g. See Section 2.2 for a discussion of the accuracy of OPIS stage data
h. Chemotherapy use collected as a binary variable (i.e. yes/no) with no information on specific drugs.

over time as a predictor of survival, probably because of the somewhat inexact staging data SEER contains. Generally these studies compare stage distributions in populations of differing ages,[85-88] ethnicities,[87,89-91] or histologic type.[92,93] Only one study examines these stage differences and links them to mammography records.[88] The single paper with significant information about chemotherapy[86] did not analyze information about specific drugs, and looked at differences in chemotherapy utilization across age groups, but not over time. It did not model survival data with systemic therapy use as one of the covariates.

Seven population-based survival analyses from Canadian authors have been published. One from Ontario included women with node-negative breast cancer and compared the survival of those who were initially treated at a teaching hospital to those who were treated at community hospitals[94]. This study did not look at the use of systemic therapy in this population. The second population-based study from Ontario exclusively examined the use of surgery and radiation therapy for the treatment of EBC between 1982 and 1992.[95] It used mastectomy-free overall survival as its primary outcome because variability in these interventions has not consistently been shown to be associated with differences in overall survival.

A 1999 paper from British Columbia, which compared the survival of breast cancer patients who were regular attendees of the provincial screening program to those who were not, included some data on the utilization of systemic therapy in this population but did not examine any trends over time.[25]

The fourth Canadian population-based survival analysis has been published in abstract form only.[96] This study from British Columbia looked exclusively at women

with metastatic breast cancer (MBC) and compared their survival over several time intervals. These time intervals were defined by the approval of new drugs for MBC, and the study linked improvements in overall survival to the availability of new drugs. However, information about the use of specific drugs in individual patients was not collected.

Another study, published only in abstract,[97] compared the outcomes of Quebec women treated according to practice guidelines to those treated in a manner other than that specified by guidelines. In the data presented to date the authors have not clarified what database was being used to define the study population, the sample size, or what specific information on patients and therapies was collected.

The most significant Canadian paper[98] is studies a question similar to those in this thesis. This paper, published in 1994 by Olivotto et al, studied the effect of adjuvant systemic therapy on breast cancer survival in British Columbia.[98] The authors examined all patients in British Columbia diagnosed with breast cancer in 1974, 1980, and 1984. This paper found an absolute 10% increase in survival for women below the age of 50 and an 8% increase in survival for women above the age of 50 over the time period studied. They concluded that these improvements in survival were related to the introduction of adjuvant systemic therapy. However, they did not collect data on stage of cancer at diagnosis, nor did they collect information about chemotherapy use at the individual patient level.

The final Canadian paper which requires comment is the analysis by Ugnat et al[99] which overlaps significantly with this thesis. Ugnat's paper was published after this thesis was approved. It analyzed survival variation with stage, age, histology, and type of

treatment using the same database as this thesis. The authors did not collect data from any patient charts[100] and restricted their analysis to patients diagnosed in Ottawa between 1994 and 1997 due to limitations in the electronic database outside of these years.

Although their analysis may be mathematically correct, it has several idiosyncrasies that makes its clinical relevance questionable. For instance, the authors lump patients with stage III (locally advanced) and IV (metastatic) breast cancer together despite the fact that these stages have substantially different prognoses and are treated entirely differently. This leads the authors to the surprising conclusion that "surgery combined with chemotherapy" leads to the best survival in stage IV disease, despite the fact that patients with metastases rarely have surgery as part of their primary treatment plan. This group had similar problems with an analysis of lung cancer patients[101] which was criticized as "not correctly presented". [102] Thus, although Ugnat's paper uses the same source and analytic methods as this study, it is felt that it is misguided in its conclusions and a continuation with this thesis is worthwhile.

1.5 Propensity Score Analysis

Retrospective review of patient data derived from standard clinical practice may incorporate bias related to the non-random assignment of patients to various treatments. For this reason it is desirable to divide patients into similar groups before comparing their outcomes. The propensity score method is one approach to achieving this. A propensity score is the predicted probability of being in the treatment (versus control) group and is determined by incorporation of baseline covariates into a multivariable logistic regression model. The propensity score reduces each patient's set of covariates to a single number,

making it possible to match or stratify on multiple variables simultaneously. Guidelines for properly generating propensity scores have been published.[107]

Having determined propensity scores for every individual in the study, several analytic options are available.[108] Patients may be stratified into groups (usually quintiles) by their propensity scores.[109] Cases and controls may be matched by one of several methods[110] and analyzed as matched pairs.

Propensity scores have been used in a few published analyses of cancer outcomes[111-113] and cancer economics.[114] Although they have desirable characteristics, as described above, a recent review suggests that propensity score analysis gives similar results to conventional multivariable regression modeling, though noting that many studies used propensity scores poorly.[115] The present thesis will use propensity scores to both stratify and match patients, and will compare the results to standard multivariable Cox proportional hazards modeling.

2.0 Materials and Methods

2.1 The Ottawa Hospital Regional Cancer Centre

The Ottawa Hospital Regional Cancer Centre (OHRCC) was known as the Ottawa Regional Cancer Centre until 2004 when it merged with the Ottawa Hospital. The OHRCC operates out of two campuses, located at the Ottawa Hospital General Site and Ottawa Hospital Civic Site. It is the sole cancer centre for the city of Ottawa and for the Eastern Ontario region. Its catchments area includes 1.2 million people. The OHRCC has over 114 000 individual patients in its records, and saw 5231 new patients in 2005. Recent years have seen an average of 800 new consults annually for breast cancer.

The OHRCC is the only available location for cancer patients in this catchments area to receive radiation therapy. Although patients could conceivably receive chemotherapy elsewhere, chemotherapeutic drugs are very expensive and are only reimbursed by the provincial government if administered in a centre affiliated with a cancer centre. During the years covered by this study there was only one medical oncologist who practiced part-time in a private office and part time at the OHRCC. All other medical oncologists in the region practiced out of the OHRCC. Therefore, the OHRCC has a near-monopoly on patients treated in its catchments area, and it is likely that the records of the OHRCC contain information on virtually all patients treated by an oncologist in the region.

2.2 Oncology Patient Information System Database

The Oncology Patient Information System (OPIS) database is the central computer system at the OHRCC. It forms the backbone of the OHRCC's appointment

booking, treatment booking, and chart tracking functions. Every patient visit at the OHRCC since 1981 is recorded, classified, and attached to a specific diagnosis. Transcriptions of all physician progress notes have been entered since 1994. A central patient registry which contains demographic and diagnostic information is complete as far back as 1960 and contains incomplete records for the time prior to this, with the oldest record being of a new patient first seen in 1935. Although largely administrative, the database is also used for research, funding, and financial planning purposes.[116]

In addition to appointment information, OPIS contains demographic information on each patient, including date of birth, address, insurance information, hospitalizations, allergies, other (non-OHRCC) physicians, and language of preference. This information is entered by clerks in the New Patient Registry and updated by Health Information Systems (HIS) staff both at the OHRCC and at the Ottawa Hospital.[116]

Information on cancer stage, grade, and type, as well as dates of recurrence have been entered with varying degrees of fidelity over the past years, depending on whether the responsibility for this data rested with clerks or physicians. Staging data is thought to be almost complete since 1997.[116]

Patient's vital status is updated in several ways. The default position is to assume that the patient is alive. HIS staff are notified if patients miss a scheduled appointment because they are deceased. They are also notified of deaths of OHRCC patients who are admitted to the Ottawa Hospital. HIS staff also review local obituaries daily to identify OHRCC patients. Finally, OPIS is linked semi-annually with the Ontario Mortality Database (OMD). This registry is administered by the provincial government, and includes information gathered from all death certificates in Ontario. Due to legal

reporting requirements, registration of deaths within Ontario is considered to be virtually complete. Via reciprocal agreements with other jurisdictions, this register also reflects the deaths of Ontario residents that occur in most other provinces and some of the United States.[117] Therefore, the only patients who are likely not to have their deaths reflected in the OPIS database are those who are not Ontario residents, or the rare cases of Ontario residents dying in distant jurisdictions (outside North America) or dying in a way that is not registered in the OMD (most commonly cases of death where no body can be identified). In 2005, 2491 OHRCC patients are known to have died, with 2274 of these deaths ascertained through local means and the remaining 217 through linkage with the OMD.

There are two published studies assessing the accuracy of the information in the OPIS database. The first looks specifically at the capture of information on tumour stage in the years 1994-1998. This study found that the staging information which was present was very accurate (incorrect in fewer than 6% of cases) but that many cases had no staging data (data present for only 71.5% of all cases reviewed).[118] The second study is disturbing on the surface because it suggests that a fraction of patients seen at the OHRCC are never registered in OPIS.[119] However, personal contact with the author of this paper alleviates some concerns. This paper is meant to be a review of cases seen by one physician, but some of his cases were entered into the system in the name of other physicians because he was either seeing them in other physicians' clinic slots, or other physicians were working with him in his clinic. Thus, all patients were entered into OPIS, but were not necessarily retrieved when he requested information on all patients

seen by him.[120] Both of these studies found that the demographic and survival data included in OPIS was close to 100% accurate.

2.3 Study Database

The Study Database consists of data retrieved directly from the OPIS database and data collected by hand from review of patient charts. Depending on the year of diagnosis charts exist as electronic documents, conventional paper charts, or as microfilm. Every patient included in the study database had their chart manually reviewed. The initial study database included all patients referred to the OHRCC with a new diagnosis of breast cancer in the calendar years 1985, 1988, 1992, 1995, 1998, and 2001. This database was compiled by HIS technicians at the OHRCC who used the date registered as the "New Disease Date"

The variables collected, and their sources, are described in table 2.1. From these variables several other variables were generated which also appear in the study database. These are described in table 2.2.

The database used for analysis includes Study ID as the sole unique identifier: patient name, OHRCC chart number, and Health Insurance Number were removed from the database as per privacy and confidentiality guidelines established by the Ottawa Hospital Research Ethics Board.

Table 2.1. Variables in Study Database

Variable Name	Variable Type	Description	Source
Patient Name	Categorical	Patient's name	OPIS
Chart Number	Categorical	OHRCC chart number	OPIS
Birth Date	Date	Patient's date of birth	OPIS
Status	Categorical	0=dead 1=alive at last follow up	OPIS
Diagnostic Date	Date	Date of breast cancer diagnosis	OPIS
Best Site	Categorical	Description of tumour side and location within breast	OPIS
Diagnosis Number	Categorical	Number of cancer diagnoses, including breast cancer diagnosis	OPIS
Last Contact Date	Date	Date of last contact with OHRCC, or date of death for patients known to be dead	OPIS
New Disease Date	Date	Date referred to OHRCC for current breast cancer diagnosis	OPIS
Gender	Categorical	0=female 1=male	OPIS
Diagnostic Age	Continuous	Age (in years) at diagnosis	OPIS
Histo	Categorical	Histologic type of breast cancer, by ICD codes	OPIS
Diag Postal Code	Categorical	Postal code at diagnosis	OPIS
Province	Categorical	Province of residence at diagnosis	OPIS
Insurance Abbr	Categorical	Name of insurance provider	OPIS
Insurance	Categorical	Insurance number	OPIS
Year	Categorical	Year of current breast cancer diagnosis	OPIS
Tumour stage	Categorical	T stage as per 1988 TNM classification	Charts
Tumour size	Continuous	Maximum dimension of tumour in millimeters	Charts
Nodal stage	Categorical	N stage as per 1988 TNM classification	Charts
Positive nodes	Continuous	Number of nodes found to contain tumour	Charts
Total nodes	Continuous	Number of nodes removed at surgery	Charts
Overall stage	Categorical	Stage grouping as per 1988 TNM classification	Charts
ER	Categorical	0=tumour is estrogen receptor negative 1=tumour is estrogen receptor positive 2=estrogen receptor status unknown	Charts
PR	Categorical	0=tumour is progesterone receptor negative 1=tumour is progesterone receptor positive 2=progesterone receptor status unknown	Charts
Med Onc	Categorical	Initials of treating medical oncologist	Charts
Rad Onc	Categorical	Initials of treating radiation oncologist	Charts
Neo	Categorical	0= No systemic therapy, or systemic therapy administrated adjuvantly 1= Neoadjuvant systemic therapy	Charts
Chemo	Categorical	Chemotherapy regimen administered	Charts
Tamoxifen	Categorical	0=No hormonal therapy 1=Hormonal therapy administered	Charts
Grade	Categorical	Histologic grade of tumour, when known Either 1-3 (Fisher Nuclear Grade) or 3-9 (SBR Grading system)	Charts

Table 2.2. Variables Generated for Study Database			
Variable Name	**Variable Type**	**Description**	**Source Variables**
Study ID	Categorical	Individual patient identifier generated for the study	
Binary Chemo	Categorical	0=no cytotoxic adjuvant chemotherapy 1=cytotoxic adjuvant chemotherapy	Chemo
Anthracycline	Categorical	0=no anthracycline-based adjuvant chemotherapy 1= anthracycline-based adjuvant chemotherapy	Chemo
Low grade	Categorical	0=Not low grade, or unknown grade 1=Known low grade (FNG=1, SBR=3-5, or pathologic description)	Grade
High Grade	Categorical	0=not high grade, or unknown grade 1=Known high grade (FNG=3, SBR=8-9, or pathologic description)	Grade
Tri-Grade	Categorical	1= Known low grade (FNG=1, SBR=3-5, or pathologic description) 2=Intermediate or unknown grade 3= Known high grade (FNG=3, SBR=8-9, or pathologic description)	Grade
Nodal	Categorical	0=Nodes negative 1=Any nodes positive	Nodal Stage
Systemic	Categorical	0=No systemic adjuvant therapy of any type 1=Some systemic adjuvant therapy	Chemo, Tamoxifen
Time	Continuous	Time between diagnosis and death or censoring	Diagnostic date, Last Contact Date
Tumour Size	Categorical	1=Right-sided breast tumour 2=Left-sided breast tumour	Best Site

2.4 Inclusion Criteria

The study database included all patients seen at the OHRCC with a new diagnosis of breast cancer during the years 1985, 1988, 1992, 1995, 1998, and 2001. No patients were removed from the database unless they met one of the exclusion criteria, defined below.

2.5 Exclusion Criteria

Patients were excluded from the study if they met any of a series of pre-determined criteria. These criteria, with a brief rationale for each, are:

Non-resident of Ontario

The vital status of patients in the OPIS system is updated by several means, including periodic linkage with the Ontario Mortality Database. Patients who are treated at the OHRCC but live in Quebec or other provinces may not have their vital status updated in OPIS. These patients are excluded as this study requires accurate information about vital status and date of death.

Metastatic breast cancer at presentation

This study focuses on outcomes related to adjuvant systemic therapy. By definition, patients with metastatic disease cannot receive adjuvant therapy. Patients who develop metastatic disease after their initial diagnosis and treatment were not excluded from the database. According to the 1988 TNM staging system, patients are considered to have metastatic disease if there is disease anywhere other than the breast, ipsilateral axillary lymph nodes, or ipsilateral internal mammary nodes. Tumour in the skin over the breast or the chest wall beneath it are not considered metastatic disease if the tumour is contiguous with the tumour in the breast. Tumours in supraclavicular lymph nodes were considered metastatic disease in the 1988 system.

Referred for a breast cancer diagnosed or treated in an earlier calendar year

Patients who move to Ottawa from elsewhere are often referred to the OHRCC for continued follow-up of a cancer diagnosed in a previous year. Alternately, they may

be referred with a recurrence of a previously diagnosed and treated cancer. These patients should not be considered in the same group as patients with a new diagnosis of breast cancer.

ORCC chart lost or irretrievable

Some charts, particularly in older cohorts, simply could not be located. These patients were excluded unless there was some other way of obtaining accurate staging and treatment information.

Patient referred for diagnosis other than carcinoma of the breast

Some patients in the database did not have carcinoma of the breast. In some cases patients had other breast diagnoses such as sarcomas of the breast or carcinoma *in situ*. Others were simply misclassified, clearly having presented with some other type of cancer.

2.6 Accuracy of Study Database

For purposes of assessing the accuracy of the study database, the paper chart was considered the gold standard. Fifty randomly selected charts were reviewed for discrepancies between them and the database. Every variable in the study database was checked against the paper chart to assess both the accuracy of information in the OPIS database and the accuracy of information abstracted manually from charts.

Missing Data

There are virtually no empty fields in the data generated from the OPIS database; for these variables it is more important to assess their accuracy than their completeness. Methods to assess the accuracy of the database are described above.

Variables collected by manual review of the chart were subject to two types of missing data: those where results for individual variables could not be found, and those where the chart could not be located. Patients in the latter group were excluded from the study because all variables were missing

For patients where individual values could not be determined from the chart the mean value of that variable for that cohort was imputed. If the variable was described by its TNM value, but not by the continuous underlying variable (e.g. tumour described as "T1c" but the exact size not mentioned), then the mean value of that variable for that cohort and TNM subset was imputed. The status of estrogen receptors, progesterone receptors, and grade of the tumour are occasionally unknown in clinical practice. In these cases the variables were coded as unknown, rather than imputing a value.

2.7 Methods of Analysis

2.7.1 Description of Data

Categorical variables are described in contingency tables. Where appropriate, their relationships are examined with χ^2 or Fisher's Exact tests. Continuous variables are described by their mean and standard deviation if approximately normally distributed, and by their median and range if their distribution is skewed.

2.7.2 Generation of Propensity Score

A propensity score is assigned to each individual in the study and represents the conditional probability of that individual receiving the study treatment. The score is determined by logistic regression using each patient's baseline covariates, as well as appropriate interaction terms. Logistic regression was carried out with SAS version 8.0.

The association between categorical variables and administration of chemotherapy was assessed using χ^2 or Fisher's Exact tests.

Statistical and graphical methods were used to identify the most appropriate method for associating the continuous variables with chemotherapy use. Univariate logistic regression was carried out with continuous variables both in their untransformed form, and divided into quintiles. If the difference in the –2logL scores of these models exceeded the χ^2 value with 3 degrees of freedom the association of the variable with chemotherapy use was deemed non-linear.

For variables requiring transformation, a smoothed curve of the variable versus predicted probability of chemo administration was examined. The shape of this curve was used to determine appropriate transformations of these variables. The transformed forms were then compared with the ranked quintiles by the same procedure described above.

Once all relevant variables were identified and appropriately transformed for analysis multivariable logistic regression was carried out. An initial model was generated with all variables whose univariate association with chemotherapy administration had a p-value of less than 0.25. Variables were then removed one at a time in descending order of p-value. After each variable was removed, the covariates of the remaining variables were examined for changes of greater than ten percent, which might indicate confounding with the removed variable.

Interaction terms were then considered. All first-order interactions between variables in the model were examined. The interaction terms were added to the model by

a process of automatic forward selection using SAS default entry and exit (0.05 for both). The fit of the final model was assessed with the Hosmer and Lemeshow Goodness of Fit test. The predictive adequacy of the model was assessed with the c-statistic (see section 5.1)

2.7.3 Survival Analysis Stratified by Propensity Scores

The propensity scores were used to stratify patients into quintiles. The adequacy of the propensity score for generating strata of similar patients was assessed within each stratum by means of chi-squared or Fisher's Exact tests for categorical variables, and t-tests for continuous variables. Survival analysis was then carried out to determine the effect of systemic chemotherapy on overall survival within each stratum. This was accomplished with simple Kaplan-Meier curves and log-rank tests for strata with no imbalanced variables. For strata with imbalanced variables Cox proportional hazards modeling was carried out with adjustment for the imbalanced variables only.

2.7.4 Survival Analysis of Populations Matched by Propensity Scores

Two data sets were created: one contained patients who had received chemotherapy (cases), the other contained patients who had not (controls). Potential matches are those where the difference in the logit of the treated and untreated patient are less than a value, termed the caliper. As per the method of Martin and Ganguly[121] the value of the caliper is

$$0.2 * \sqrt{[(VARcas+VARcon)/2]}$$

where VARcas and VARcon are the variance of the logit of case and control groups, respectively. If there were multiple potential controls for any case then these were ranked from best match (smallest difference in logits) to worst match. Cases and their best match were then placed in the final data set in random order. If the best match for a case was removed because it was also the best match for a previously selected case then the best remaining match was taken. If all possible matches for a particular case were taken by other cases then that case was not placed in the final data set.

The characteristics of the case and control groups were compared with χ^2 or Fisher's Exact tests for categorical variables and t-tests for continuous variables. A Kaplan-Meier curve and log-rank test were then performed to compare overall survival in the two groups. Paired analysis of survival was undertaken with conditional logistic regression, a variant of logistic regression appropriate for analysis of paired samples.

2.7.5 Conventional Survival Analysis

To better assess the value of propensity score analyses compared to analyses not using propensity scores, a conventional Cox proportional hazards analysis was undertaken. Each variable was first assessed for its relationship to survival in a univariate setting. The assumption of proportional hazards was tested by evaluation of a plot of log-log survival versus log time for categorical variables.

To determine the best format for continuous variables, Martingale residuals for a null model were calculated and plotted against values of the variable, with a smoothed curve superimposed to aid in interpretation. A straight line indicates an appropriate

functional form of the variable.[122] The curve was smoothed using the loess procedure, a moving linear regression procedure.[123] (See section 5.1)

Once all relevant variables were identified and appropriately transformed for analysis multivariable Cox proportional hazards regression was carried out. An initial model was generated with all variables whose univariate association with chemotherapy administration had a p-value of less than 0.25. Variables were then removed one at a time in descending order of p-value. After each variable was removed, the covariates of the remaining variables were examined for changes of greater than ten percent, which might indicate confounding with the removed variable.

Interaction terms were then considered. All first-order interactions between variables in the model were examined sequentially. Interaction terms were retained if they were statistically significantly associated with survival or if they altered the covariate of their component variables by more than 10%.

3.0 Results

3.1 Description of the Study Database

During the study years, 4095 patients were registered in OPIS with a new diagnosis of breast cancer. Of these, 470 were excluded because they were not Ontario residents. The charts of the remaining 3625 patients were retrieved and reviewed. 562 patients were excluded for medical reasons including the presence of metastatic disease at presentation, recurrent disease, or having a neoplastic diagnosis other than breast carcinoma. Also, 78 patients were excluded for administrative reasons including the inability to retrieve the chart, or because they appeared more than once in the database as duplicate records. There were a large number (43% of total number) of irretrievable charts for patients treated at the General Campus in the 1985 cohort. It was therefore decided to exclude every patient treated at the General Campus during that year, since the inclusion of only the retrievable charts may introduce selection bias. These patients are not included in the original total of 4095 patients.

This left a total of 2985 patients eligible for analysis. Patient numbers and reasons for exclusion are presented in table 3.1.

The characteristics of all included patients along with their tumours are presented in tables 3.2-3.5. Data is given for each cohort year of the study.

Several issues are notable in these tables. Firstly, the proportion of lost charts is greater in the in the earlier cohorts with a progressive decline over time (Table 3.1). The OHRCC switched to electric charts in 1994, which explains why no charts are lost after

Table 3.1. Reasons for Exclusion from Study							
	1985*	1988	1992	1995	1998	2001	Total
Patients in Database	381	574	771	768	785	813	4092
Lost Charts	34	16	15	0	0	0	65
Non-Ontario Resident	38	98	133	104	37	60	470
Stage IV Disease**	25	38	43	36	55	63	260
Previous Diagnosis***	44	36	44	54	42	32	252
DCIS/LCIS/Paget's Only†	2	4	8	7	6	5	32
Not Breast Cancer	1	2	0	2	2	0	7
Phylloides Tumour††	0	2	1	2	2	1	8
Duplicate Records	1	1	2	2	2	1	9
Other	0	2	2	0	0	0	4
Total Eligible	**236**	**375**	**523**	**561**	**639**	**651**	**2985**

* The 1985 cohort includes only patients treated at the Civic site. The General site is excluded because of a large number (43%) of irretrievable charts from this year

** Ontario patients with metastatic disease at the time of presentation

*** Patients seen in consultation at the OHRCC in order to continue follow-up for a cancer diagnosed in an earlier year, or for treatment of a recurrence of a cancer diagnosed in an earlier year

† Ductal carcinoma in situ (DCIS), lobular carcinoma in situ (LCIS) and Paget's disease of the nipple are conditions of the breast which are occasionally managed at the cancer centre, though they are not invasive breast cancers.

†† Also known as cystosarcoma phylloides, this is a sarcoma of the breast. These patients are excluded because the treatment and natural history of this disease differs substantially from carcinoma of the breast.

this date. Pre-1994 charts exist in paper form only and most are stored in various warehouses around Ottawa. Over the years, some charts have been misfiled, damaged, or simply lost.

The second significant issue is a decrease in tumour size over time (Table 3.4). With the exception of 2001, the mean size of tumours decreased each year. Between 1985 and 1998, the mean tumour size decreased from 28.12 mm to 21.86 mm (p-value for trend <0.0001). The mean tumour size in 2001 increased to 26.02 mm. Without data from subsequent years it cannot be ascertained whether this represents a single aberrant year or the start of a new trend in tumour size over time.

Table 3.2. Stage Distribution of Patients by Year							
	1985	1988	1992	1995	1998	2001	Total
Overall Stage							
Stage1	75	134	230	253	331	292	1315
Stage 2A	86	125	168	171	175	176	901
Stage 2B	53	78	64	86	96	114	491
Stage 3A	10	19	31	23	18	40	141
Stage 3B	12	19	30	28	19	29	137
T-stage							
T1a	3	11	24	38	31	27	134
T1b	21	39	76	93	119	94	442
T1c	77	136	192	189	258	238	1090
T2	111	142	169	185	183	211	1001
T3	12	28	33	28	29	54	184
T4a	2	5	4	3	5	8	27
T4b	5	9	9	10	4	10	47
T4c	3	3	3	2	2	0	13
T4d	2	2	13	13	8	9	47
N-stage							
N0	136	216	318	325	401	389	1785
N1a	1	6	5	9	12	11	44
N1bi	53	79	80	93	78	98	481
N1bii	24	21	23	24	35	38	165
N1biii	4	15	31	34	36	48	168
N1biv	6	16	17	16	28	21	104
N2	3	9	10	8	3	8	41
N3	1	1	1	0	0	2	5
Unknown	8	12	38	52	46	36	192

A third significant trend is the increasing likelihood over time that patients are treated with chemotherapy. The proportion of patients receiving chemotherapy increases by an absolute 1.8% during each year of the study period (χ^2 test for trend: p =0.0004). In 1985, 17.4% of patients received chemotherapy. In 2001, this number was 47.5%. A prominent component of this increase occurred in the node-negative population, which saw a ten-fold increase in use of chemotherapy from 2.8% in 1985 to 29.4% in 2001. The node-positive population saw a smaller relative increase in chemotherapy use, from 40.2% in 1985 to 81.4% by 2001.

Table 3.3. Other Categorical Tumour Characteristic by Year							
	1985	1988	1992	1995	1998	2001	Total
ER Positive	134	233	352	330	442	493	1984
ER Negative	66	70	57	141	139	140	613
ER Unknown	36	72	114	90	58	18	388
PR Positive	117	223	300	313	373	451	1777
PR Negative	74	80	108	158	208	182	810
PR Unknown	45	72	115	90	58	18	398
Low Grade	6	32	112	104	191	209	654
Unknown or Intermediate Grade	211	298	301	356	329	270	1765
High Grade	19	45	110	101	119	172	566

Table 3.4. Continuous Variables by Year								
		1985	1988	1992	1995	1998	2001	Total
Tumour Size (mm)	Mean	28.12	26.58	25.24	24.36	21.86	26.02	24.92
	St. Dev	16.77	17.95	20.78	20.83	18.35	21.74	19.95
Number of Positive Nodes	Mean	1.40	1.28	1.09	0.99	1.02	1.26	1.15
	St. Dev	2.47	2.25	2.17	1.95	2.10	2.45	2.22
Age at Diagnosis (years)	Mean	58.64	58.45	59.41	59.32	61.15	59.42	59.59
	St. Dev	14.00	13.24	13.00	13.96	13.51	13.43	13.52
Follow-Up Time (years)	Mean	9.55	9.33	8.00	5.83	3.95	2.58	5.88
	Maximum	19.91	16.50	13.94	8.82	5.82	3.53	19.91

Table 3.5. Systemic Therapy by Year							
	1985	1988	1992	1995	1998	2001	Total
No Systemic therapy	175	206	209	185	172	115	1062
Chemotherapy Only, Non-anthracycline	30	33	42	24	8	2	139
Chemotherapy Only, Anthracycline	3	29	44	84	72	108	340
Hormonal Therapy Only	20	87	174	185	232	227	925
Chemotherapy and Hormonal Therapy Non-Anthracycline	8	13	16	15	10	4	66
Chemotherapy and Hormonal Therapy, Anthracycline	0	7	38	68	145	195	453
Total	236	375	523	561	639	651	2985

3.1.1 Accuracy of the study database

A total of 50 charts were obtained and all variables in the study database were compared with the values reported in the chart. This was undertaken to assess the accuracy of the electronic information in OPIS. The paper charts were considered the gold standard. The results can be found in table 3.6.

Table 3.6. Accuracy of information in electronic database

Variable	Correct	Incorrect	Absent
Date of Birth	100%	0%	0%
New Disease Date	100%	0%	0%
Diagnostic Date	98%	2%	0%
Last Contact Date	96%	4%	0%
Vital Status	98%	2%	0%
Age at Diagnosis	100%	0%	0%
Postal Code at Diagnosis	88%	8%	4%
T-stage	14%	50%	36%
N-stage	24%	30%	46%
M-Stage	54%	4%	42%
Total Stage	20%	34%	46%
Chemotherapy administered	94%	6%	0%
Tamoxifen administered	84%	16%	0%
Side of tumour	100%	0%	0%
Histologic subtype of tumour	100%	0%	0%
Gender of patient	100%	0%	0%

Overall, the data in OPIS is very accurate with the exception of staging data and, to a lesser extent, treatment data. For this reason, all staging and treatment data for this thesis were collected by hand from the charts, rather than from OPIS.

There are two charts with erroneous last contact dates, and these merit further discussion because of the centrality of this variable to the analysis. The first error was typographical, being the transposition of digits such that a patient was recorded as having been last seen on the seventh day of the sixth month, rather than the sixth day of the seventh month. The second error involved a woman with an exceptionally common

name (there are 10 women with the same name in the OHRCC database) who was recorded as having died on a particular date, when in fact the death had happened to another woman of the same name.

3.2 Generation of Propensity Score for All Patients

Note that those tables that present results are included within the text, while those charts and graphs which do not contain results but were essential for the generation of models are included in the section titled 'Supplementary Tables'.

3.2.1 Examination of the Variables

We determined which variables were associated with administration of chemotherapy using χ^2 or Fisher's Exact tests for categorical variables (Supp. table 3.1s). The only variable with no association was the side of the tumour. The identity of the medical oncologist was also excluded from consideration because of the large number of oncologists who treated a small number of cases. Of twenty-four oncologists identified, eleven had treated fewer than twenty cases.

The suitability of variables for inclusion in the multivariate model was assessed by the method outlined in section 2.7.2. Continuous variables were evaluated to determine whether they could be used in untransformed or transformed formats. Patient age at diagnosis required no transformation (Supp table 3.2s, Supp graph 3.1s). Tumour size required natural logarithmic transformation (Supp table 3.3s, Supp graphs 3.2s, 3.3s). Number of positive nodes was transformed into a dichotomous rather than continuous variable (Supp graph 3.4s).

The continuous variable tumour size was used, rather than the categorical variable t-stage. T-stages one through three correspond directly with specific tumour sizes. T-stage four includes tumours of any size that invade the overlying skin or the chest wall beneath the tumour. To capture this information a dichotomous variable was added that took a value of zero if the tumour was T-stage one through three, and one if the T-stage was four.

3.2.2 Multivariable models

An initial model was generated with all variables described in Tables 3.1s-3.3s using the methodology described in section 2.7.2. The variables remaining in the model were grade, ER status, nodal status, year of diagnosis, age at diagnosis, tumour size (transformed as described above), and whether the tumour was T-stage four.

Interaction terms were then considered. This resulted in six interaction terms being added to the model. Four of these terms included the nodal status, combined with grade, year of diagnosis, age at diagnosis, and tumour size. The other interaction terms in the model were between ER status and age, and between grade and tumour size.

The maximum likelihood estimates and odds ratio estimates are found in table 3.7. The c-statistic for the model is 0.962. The Hosmer and Lemeshow test suggests a good fit of the model with the data (p=0.7369).

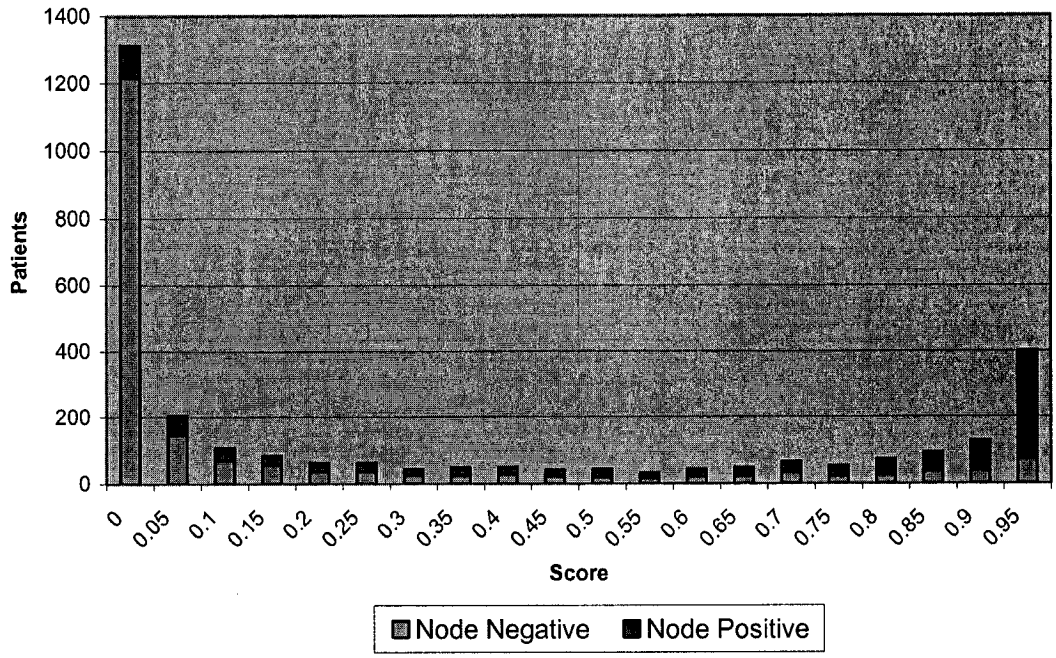
The most prominent feature of this model is the very high odds ratio for nodal status. Because this variable is a component of four interaction terms its odds ratio cannot be interpreted alone. For many of the years covered by this study it was clinical dogma that patients with positive lymph nodes required adjuvant systemic therapy, while those who were node-negative did not. Therefore it is not surprising that this variable is

Table 3.7. Propensity Score Model for All Patients						
Parameter	**Level**	**Estimate**	**Standard Error**	**p-value**	**OR Point Estimate**	**Wald 95% Confidence Limits**
Intercept		-8.6760	1.2636	<0.0001		
Grade	2 vs. 1	3.3646	0.6951	<0.0001	28.921	7.406-112.948
	3 vs. 1	6.9588	1.2860	<0.0001	>999.999	84.626->999.999
Estrogen Receptors	Pos vs. neg	-0.3314	0.4945	0.5027	0.718	0.272-1.892
	Unk vs. neg	2.6436	1.4171	.0.0621	14.064	0.875-226.128
Nodal		12.2331	1.2390	<0.0001	>999.999	>999.999->999.999
Year	1988 vs. 1985	1.1193	0.7299	0.1251	3.063	0.733-12.804
	1992 vs. 1985	3.1038	0.6597	<0.0001	22.283	6.115-81.193
	1995 vs. 1985	4.0505	0.6574	<0.0001	57.424	15.830-208-310
	1998 vs. 1985	4.4259	0.6574	<0.0001	83.592	22.665-308.296
	2001 vs. 1985	5.3557	0.6659	<0.0001	211.822	56.834-789.464
Patient age		-0.1043	0.0104	<0.0001	0.901	0.883-0.920
Ln(tumour size)		3.4788	0.4754	<0.0001	32.422	12.770-82.315
T4		1.3740	0.3311	<0.0001	3.951	2.065-7.560
Grade*nodal	1 vs. 0	2.5332	0.5061	<0.0001	12.539	4.670-33.959
	2 vs. 0	0.7923	0.3532	0.0243	2.208	1.105-4.413
Nodal*year	1985 vs. 0	1.4762	0.8031	0.0660	4.376	0.907-21.122
	1988 vs. 0	1.3183	0.5999	0.0280	3.737	1.153-12.111
	1992 vs. 0	-0.3193	0.4819	0.5016	0.727	0.283-1.869
	1995 vs. 0	-0.7948	0.4496	0.0771	0.452	0.187-1.090
	1998 vs. 0	0.9562	0.4359	0.0283	2.602	1.107-6.114
Nodal*age		-0.0680	0.015	<0.0001	0.934	0.907-0.962
Nodal*ln(size)		-1.9576	0.2499	<0.0001	0.141	0.087-0.230
Grade*ln(size)		-0.6571	0.2009	0.0011	0.518	0.350-0.768
ER status*age		-0.0225	0.00805	0.0052	0.978	0.962-0.993

highly predictive for the receipt of adjuvant chemotherapy. In the model without interactions the odds ratio associated with node-positivity was 32. The drastic increase with the addition of interaction terms suggests that there is quasi-complete separation of the data around this variable (See section 5.1). Although a stratified survival analysis was carried out with the full model, it was decided to create separate models for node-positive and node-negative patients, and to repeat the analyses on these separate models. The process for creating these models is described below.

The propensity scores have a striking bimodal distribution (Graph 3.1). Off all scores, 42% are less than 0.025, and 13% are greater than 0.975. For patients who are node-negative, the median propensity score is 0.029, while it is 0.82 for patients who are node-positive.

Figure 3.1 Distribution of Propensity Scores



3.2.3 Adequacy of the model for creating comparable groups

Propensity score adequacy was measured by dividing the patients into quintiles of propensity score and comparing baseline variables by chemotherapy status within each quintile.[124]

For the model including all patients, the results by quintile can be found in tables 3.8-3.9. There is no table for the first quintile (the quintile least likely to receive chemo) because only one patient in this quintile received chemotherapy. In each of the quintiles, one or two of the seven variables is significantly different between the chemotherapy and non-chemotherapy arms. I shall refer to these variables as "imbalanced". In several cases, such as the imbalance in Year of Treatment in quartile two, or Nodal Status in quartile five, the imbalance arises from a small number of patients in one or the other of the arms. The imbalances that seem most clinically significant are the difference in mean

tumour size between chemo and non-chemo arms in quintile three (24.98mm vs. 20.85 mm) and the younger age of chemo patients in quintile four (55.26 years vs. 57.52 years). Had the survival analysis been carried out without any attempt at adjusting for imbalances between chemo and non-chemo groups, then every single variable in the database would have been imbalanced, except for the side of the tumour (data not shown).

Table 3.8. Comparability of Chemo and Non-Chemo Groups within Propensity Score Quintiles for All Patients. Continuous Variables.

Quintile	Variable	Chemo		No Chemo		
		Mean	St. Dev	Mean	St. Dev	p-value
2/5	Age at Diagnosis	61.375	13.341	64.657	11.67	0.4306
	ln(tumour size)	3.0118	0.5717	2.7459	0.6274	0.2338
3/5	Age at Diagnosis	62.056	11.811	61.388	12.118	0.6618
	ln(tumour size)	3.2189	0.6326	3.0375	0.5851	0.0154
4/5	Age at Diagnosis	55.261	10.228	57.524	12.565	0.0158
	ln(tumour size)	3.207	0.6071	3.1994	0.5771	0.8778
5/5	Age at Diagnosis	45.732	8.6822	48.37	11.54	0.1296
	ln(tumour size)	3.3979	0.649	3.5032	0.5189	0.4069

Table 3.9. Comparability of Chemo and Non-Chemo Groups within Propensity Score Quintiles for All Patients. Categorical Variables.													
		Quintile 2/5			Quintile 3/5			Quintile 4/5			Quintile 5/5		
Variable	Level	Chemo	No Chemo	p-value	Chemo	No Chemo	p-value	Chemo	No Chemo	p-value	Chemo	No Chemo	p-value
Nodal	Negative	6	526	0.1970	47	371	0.4543	179	133	0.5728	120	11	0.0157
	Positive	2	63		24	155		170	115		450	16	
Grade	Low	1	166	0.4471	15	96	0.8370	29	26	0.3627	66	2	0.3192
	Intermed./ Unknown	7	393		46	356		203	151		254	16	
	High	0	30		10	74		117	71		250	9	
ER Status	Positive	3	52	0.0207	47	395	0.2330	223	160	0.6111	303	13	0.7755
	Negative	4	426		16	80		113	75		237	13	
	Unknown	1	111		8	51		13	13		30	1	
Year	1985	3	39	0.0195	4	41	0.7326	16	14	0.5421	18	3	0.0963
	1988	2	91		13	64		26	28		41	2	
	1992	1	91		11	103		61	44		67	2	
	1995	0	130		13	97		87	51		91	7	
	1998	1	131		14	100		79	49		141	8	
	2001	1	107		16	121		80	62		212	5	
T4	T4	0	11	0.8610	8	17	0.0015	19	14	0.9156	55	4	0.3288
	Not T4	8	578		63	509		330	234		515	23	

3.3 Generation of Propensity Score for Node-Positive Patients

3.3.1 Examination of the Variables

The study database included 1008 node-positive patients, 646 of whom received adjuvant chemotherapy. All variables of interest were examined individually for their relationship with chemotherapy administration, using the method described above. These results are summarized in Supp table 3.4s. Tumour side had no association with chemotherapy administration. The relationship between tumour size and treatment with chemotherapy was weaker in this population than in the whole patient population.

The appropriate format for continuous variables in the node-positive model was determined using the process described above (section 2.7.2). Age at diagnosis did not require transformation (see Supp table 3.5s and Supp graph 3.5s). Tumour size and the number of positive nodes could not be made linear with simple transformations due to their non-monotonic forms (see Supp tables 3.6s-3.7s and Supp graphs 3.6s-3.7s). These variables were therefore divided into quintiles and analyzed as categorical variables. Because most patients had either one or two positive nodes, node number was divided into quartiles.

3.3.2 Multivariable models

An initial model was generated with all variables described in Tables 3.4s-3.7s. The model was generated by the method described above. The variables included in the final model are year of diagnosis, age at diagnosis, number of nodes positive, and estrogen receptor status. No significant interactions were observed between any of the variables.

The maximum likelihood estimates and odds ratio estimates are found in table 3.10. The c-statistic for the model is 0.940. The Hosmer and Lemeshow goodness of fit test suggests an adequate fit of the model to the data (p=0.0735)

The distribution of propensity scores is still bimodal, with 10% of scores being less than 0.025, and thirty four percent greater than 0.975. The median score is 0.81, which is very similar to the median score for the node-positive subgroup of the model including all patients.

Table 3.10. Propensity Score Model for Node-Positive Patients

Parameter	Level	Estimate	Standard Error	p-value	OR Point Estimate	Wald 95% Confidence Limits
Intercept		10.2788	0.7916	<0.0001		
Year	1988 vs. 1985	0.9808	0.4004	0.0143	2.667	1.216-5.845
	1992 vs. 1985	1.3855	0.3946	0.0004	3.997	1.844-8.661
	1995 vs. 1985	1.8972	0.3952	<0.0001	6.667	3.073-14.466
	1998 vs. 1985	3.9255	0.4377	<0.0001	50.678	21.490-119.508
	2001 vs. 1985	4.0306	0.4441	<0.0001	56.293	23.576-134.412
Estrogen	Pos vs. Neg	-1.4566	0.2956	<0.0001	0.233	0.131-0.416
Receptors	Unk vs. neg	-1.0138	0.4782	0.340	0.363	0.142-0.926
Patient Age		-0.1870	0.0123	<0.0001	0.829	0.810-0.850
Nodal Quartiles*	2 vs. 1	0.4281	0.2830	0.1304	1.534	0.881-2.672
	3 vs. 1	0.7513	0.2934	0.0104	2.120	1.193-3.767
	4 vs. 1	1.0016	0.2730	0.0002	2.723	1.594-4.649

* Quartile 1 = 1 node, quartile 2= 2 nodes, quartile 3=3-5 nodes, quartile 4=>5 node

Table 3.11. Comparability of Chemo and Non-Chemo Groups within Propensity Score Quintiles for Node-Positive Patients. Categorical Variables.

		Quintile 1/5			Quintile 2/5			Quintile 3/5			Quintiles 4&5/5		
Variable	**Level**	**Chemo**	**No Chemo**	**p-value**	**Chemo**	**No Chemo**	**p-value**	**Chemo**	**No Chemo**	**p-value**	**Chemo**	**No Chemo**	**p-value**
Year	1985	4	32	0.5109	9	19	0.3977	10	2	0.3168	14	2	0.0111
	1988	4	42		12	24		26	5		32	2	
	1992	1	48		8	23		42	5		40	0	
	1995	3	43		15	23		30	7		61	2	
	1998	1	10		19	19		25	11		106	1	
	2001	0	13		13	19		29	9		142	1	
Estrogen	Positive	11	160	0.7349	60	100	0.4516	125	31	0.9410	236	4	0.6510
Receptors	Negative	1	7		13	17		22	5		143	4	
	Unknown	1	21		3	10		15	3		16	0	
Nodes	1	6	89	0.2002	33	48	0.5099	54	18	0.3496	126	1	0.5490
Positive	2	1	47		10	24		39	5		71	1	
	3-4	4	23		12	26		31	7		89	3	
	5+	2	29		21	29		38	9		109	3	

3.3.3 Adequacy of the model for creating comparable groups

Propensity score adequacy was measured by dividing the propensity scores into quintiles and comparing baseline variables by chemotherapy status within each quintile.

For the model including node-positive patients, the results by quintile can be found in tables 3.11-3.12. Quintiles four and five (the quintiles most likely to receive chemotherapy) had very few patients in them who did not receive chemo and so were combined into a single group, which contains 395 treated patients and 8 non-treated. Each of the quintiles contains one imbalanced variable with the exception of quintile two, which has none.

Table 3.12. Comparability of Chemo and Non-Chemo Groups within Propensity Score Quintiles for Node-Positive Patients. Continuous Variables.

Quintile	Variable	Chemo	Mean	Standard Deviation	p-value
1/5	Age	Yes	69.462	6.4758	0.0163
		No	74.223	6.8742	
2/5	Age	Yes	64.947	8.1958	0.8033
		No	64.654	8.0803	
3/5	Age	Yes	54.833	8.3949	0.0254
		No	58.103	6.9425	
4-5/5	Age	Yes	45.152	8.4227	0.2791
		No	41.875	10.683	

3.4 Generation of Propensity Score for Node-Negative Patients

3.4.1 Examination of the Variables

The study database included 1967 node-negative patients, 342 of whom received adjuvant chemotherapy. All variables of interest were examined individually for their relationship with chemotherapy administration, using the method described above. These results are summarized in Supp table 3.8s. As expected, tumour side had no association

with chemotherapy administration. More than 40% of the medical oncologists had treated fewer than five cases, so this variable was excluded from further consideration.

The appropriate format for continuous variables in the node-positive model was determined using the process described above. Age at diagnosis did not require transformation (see Supp table 3.9s and Supp graph 3.8s). Tumour size could not be used in its untransformed format; multiple transformations were tested, with ln(size) being the best (See Supp table 3.10s).

3.4.2 Multivariable models

An initial model was generated with all variables described in Tables 3.8s-3.10s. The model was generated by the method described above. The variables included in the final model are year of diagnosis, age at diagnosis, estrogen receptor status, ln(tumour size), and histologic grade of tumour. Two interaction terms are also in the model. They are the interactions between tumour size and grade, and between tumour size and patient age.

The maximum likelihood estimates and odds ratio estimates are found in table 3.13. The c-statistic for the model is 0.957. The Hosmer and Lemeshow goodness of fit test suggests an acceptable fit of the model to the data (p=0.9117)

Table 3.13. Propensity Score Model for Node-Negative Patients						
Parameter	Level	Estimate	Standard Error	p-value	OR Point Estimate	Wald 95% Confidence Limits
Intercept		-14.5055	2.8393	<0.0001		
Year	1988 vs. 1985	1.2929	0.7718	0.0939	3.643	0.803-16.535
	1992 vs. 1985	3.4464	0.7046	<0.0001	31.387	7.888-124.891
	1995 vs. 1985	4.3724	0.7067	<0.0001	79.230	19.832-316.525
	1998 vs. 1985	4.7312	0.7143	<0.0001	113.432	27.973-459.961
	2001 vs. 1985	5.7355	0.7250	<0.0001	309.662	74.783->999.999
Estrogen	Pos vs. Neg	-1.7507	0.2147	<0.0001	0.174	0.114-0.265
Receptors	Unk vs. Neg	-1.3868	0.4130	0.0008	0.250	0.111-0.561
Ln(size)		5.8724	0.9820	<0.0001	355.097	51.814->999.999
Age		-0.0181	0.0427	0.6717	0.982	0.903-1.068
Grade	2 vs. 1	4.0234	0.9146	<0.0001	55.889	9.307-335.613
	3 vs. 1	8.3317	1.7332	<0.0001	>999.999	139.026->999.999
T4		1.4375	0.5312	0.0068	4.210	1.486-11.925
Grade*size		-0.8912	0.2747	0.0012	0.410	0.239-0.703
Age*size		-0.0334	0.0136	0.0141	0.967	0.942-0.993

This model indicates that histologic grade, tumour size, and estrogen receptor status are the major factors driving decisions about chemotherapy administration in the node-negative population, which mirrors clinical practice.

3.4.3 Adequacy of the model for creating comparable groups

Propensity score adequacy was measured by dividing the propensity scores into quintiles and comparing baseline variables by chemotherapy status within each quintile.

For the model including node-negative patients, the results by quintile can be found in tables 3.14-3.15. Quintiles one and two (the quintiles least likely to receive chemotherapy) had only one patient between them who received chemotherapy and, therefore, were not analyzed. Each of the remaining three quintiles had two imbalanced variables.

Table 3.14 Comparability of Chemo and Non-Chemo Groups within Propensity Score Quintiles for Node-Negative Patients. Categorical Variables.										
		Quintile 3			Quintile 4			Quintile 5		
Variable	Level	Chemo	No Chemo	p-value	Chemo	No Chemo	p-value	Chemo	No Chemo	p-value
Year	1985	1	17		2	14	0.9994	1	1	0.0041
	1988	2	38		5	30		0	6	
	1992	1	59		10	61		38	16	
	1995	2	90		8	55		72	28	
	1998	1	99		13	81		69	23	
	2001	1	85		17	99		107	34	
Estrogen	Positive	6	277		35	234	0.7082	103	48	0.2623
Receptors	Negative	1	36		13	73		175	56	
	Unknown	1	75		7	33		9	4	
Grade	Low	2	96		10	56	0.1844	10	6	0.0106
	Intermediate/ Unknown	6	271		40	221		101	54	
	High	0	21		5	63		176	48	
T4	T4	0	5		5	7	0.0164	18	7	0.9393
	Not T4	8	383		50	333		269	101	

Table 3.15. Comparability of Chemo and Non-Chemo Groups within Propensity Score Quintiles for Node-Negative Patients. Continuous Variables.					
Group	Variable	Chemo	Mean	Standard Deviation	p-value
3	Age	Yes	54.5	10.69	0.0455
		No	62.99	11.869	
	ln(size)	Yes	2.9614	0.5789	0.3441
		No	2.7718	0.5601	
4	Age	Yes	56.982	12.183	0.5893
		No	57.912	11.788	
	ln(size)	Yes	3.2316	0.5179	0.0066
		No	3.0171	0.5435	
5	Age	Yes	48.808	10.111	0.0025
		No	52.731	14.376	
	ln(size)	Yes	3.2855	0.5797	0.5519
		No	3.2472	0.5438	

3.5 Survival Analysis Stratified by Propensity Score

3.5.1 All Patients

The study population was divided into quintiles by propensity score and, within these quintiles, the survival of patients who received and did not receive adjuvant chemotherapy was compared by means of Kaplan Meier plots (graph 3.2) and the log rank test. For the first quintile (the quintile least likely to receive chemotherapy) there was no comparison possible because only one patient received chemotherapy. For the second and third quintiles there was no significant difference between the two groups in terms of overall survival. In quintile four the chemotherapy group had a survival advantage that almost reached statistical significance (HR 0.736 95% CI 0.540-1.002). Chemotherapy patients in quintile five had a statistically significant survival advantage (HR=0.441 95% CI 0.243-0.800). In the cases of quintiles four and five, Cox proportional hazard modeling was used to correct for variables that were imbalanced between the chemo and non-chemo groups.

These graphs show a progressively larger benefit to adjuvant chemotherapy in quintiles with higher propensity scores.

3.5.2 Node-Positive Patients

The population was analyzed by propensity score quintiles by the same method described above. Please see graph 3.3 for the Kaplan-Meier plots. Only quintile two showed no significant difference between chemo and non-chemo groups.

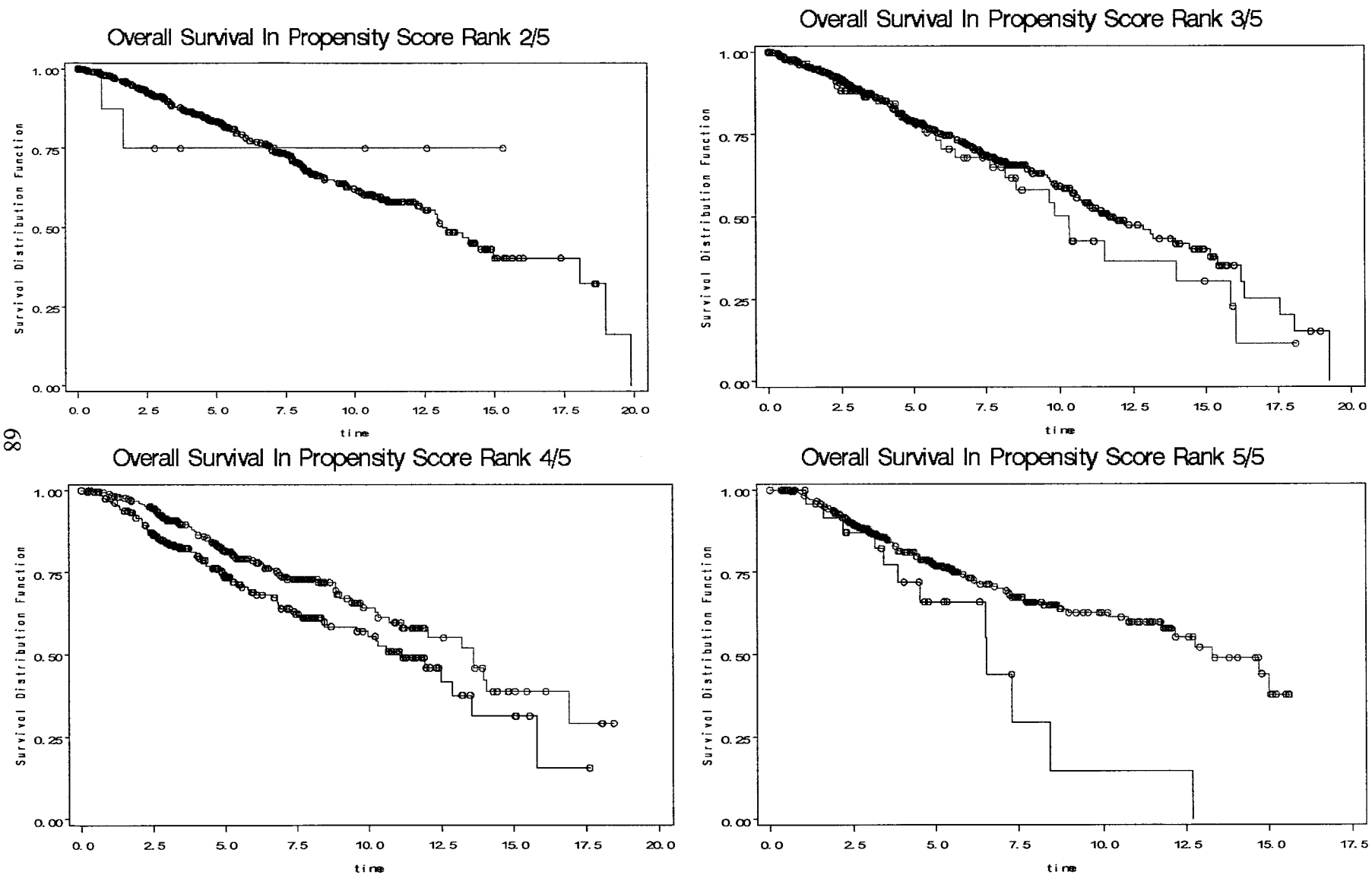


Figure 3.2. Overall survival of breast cancer patients sorted by quintiles of propensity score. Quintile 2: logrank p=0.7347. Quintile 3: logrank p=0.3118. Quintile 4: p(corrected)=0.0511 (HR 0.736 0.540-1.002). Quintile 5: p(corrected)=0.0071 (HR 0.441 0.243-0.800)

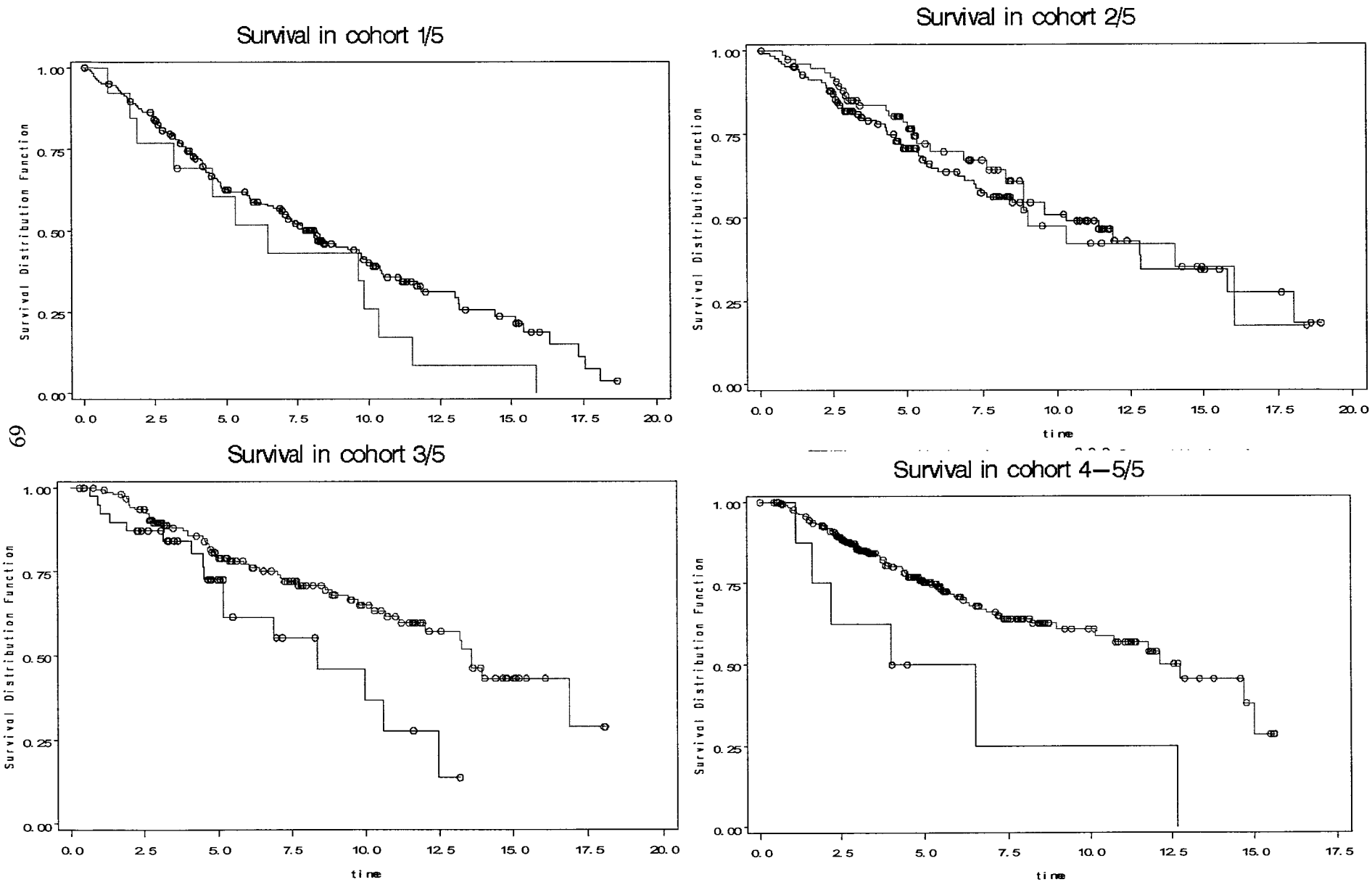


Figure 3.3 Overall survival of node-positive breast cancer patients sorted by quintiles of propensity score. Quintile 1: p(corrected)=0.0087 (HR 2.313 95% CI 1.236-4.326) Quintile 2: Logrank p=0.6028 Quintile 3: p(corrected)=0.0326 (HR 0.526 95% CI 0.291-0.948) Quintiles 4-5: p(corrected)=0.0375 (HR 0.392 95% CI 0.162-0.947)

Quintile one is notable for showing a group that seems to be harmed by adjuvant chemotherapy. Note, however, that only thirteen patients in this quintile received chemotherapy, limiting the interpretation of this result. (see Section 4.0 Discussion)

Quintiles three through five showed increasing amounts of benefit with adjuvant chemotherapy in quintiles with higher propensity scores.

3.5.3 Node-Negative Patients

The population was analyzed by propensity score quintiles by the same method described above. Please see graph 3.4 for the Kaplan-Meier plots. Note that only quintiles four and five had a sufficient number of chemotherapy-treated patients to make a meaningful analysis. There is no significant difference in overall survival in either of these quintiles, though there is a trend towards benefit in quintile five. Indeed, the uncorrected log rank test in this quintile had a p-value of 0.0120, but this difference disappeared when baseline differences in age and histologic grade were corrected for.

Thus, with this use of the propensity score, no population of node-negative patients could be found that derived an overall survival advantage from adjuvant chemotherapy.

3.6 Survival Analysis For Cases and Controls Matched by Propensity Score

3.6.1 Case-Control Matching

Cases and controls were defined and matched as per the method described in section 2.7.4. The mean propensity score in the chemotherapy patients was 0.77, while it was 0.11 in the non-chemotherapy patients. Although there were 969 cases, only 300 could be matched to appropriate controls. Although all 969 had potential matches, there

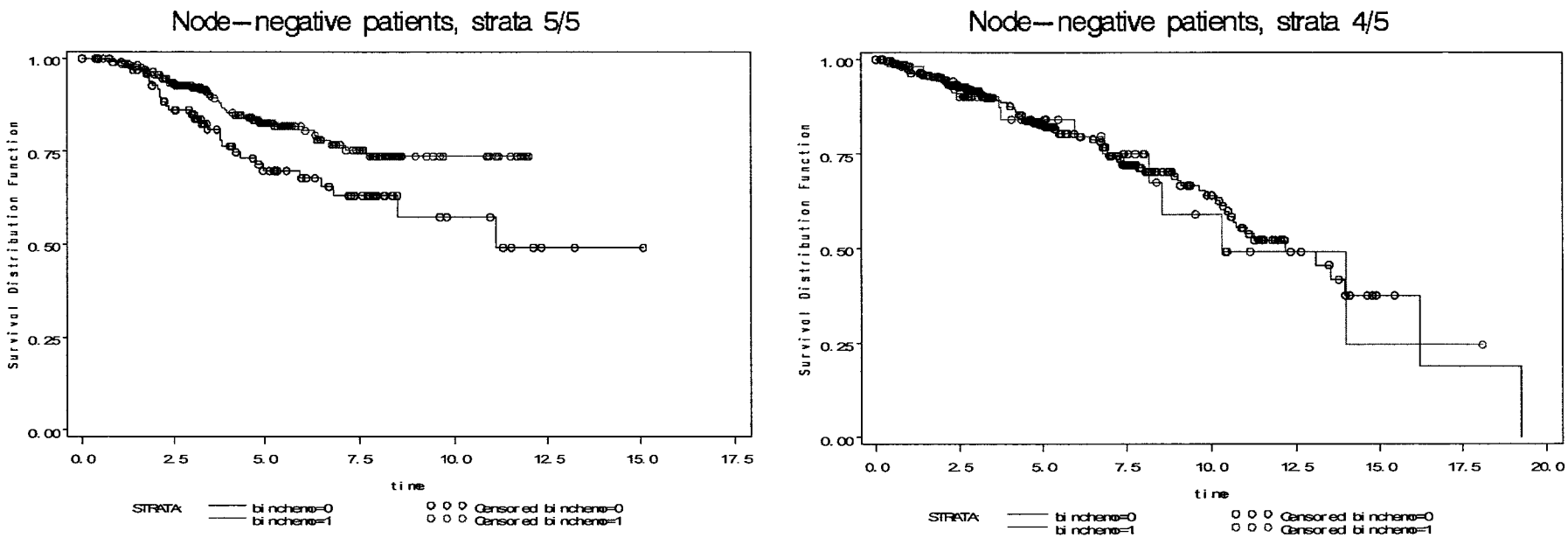


Figure 3.4 Overall survival of node-positive breast cancer patients sorted by quintiles of propensity score. Quintile 4: logrank p=0.8275 Quintile 5: p(corrected)=0.1085 (HR 0.661 95% CI 0.399-1.096) . Chemo arm in red, non-chemo arm in black.

was a relatively small number of unique controls, due to the large discrepancy in mean propensity score in the case and control groups. Thus, the removal of controls from the pool meant that many cases no longer had matches.

3.6.2 Adequacy of the matching for producing comparable groups

The case and control groups were very well matched, with no statistically significant differences in any of the measured variables (see tables 3.16-3.17). These groups are better matched than most of the groups created by stratification by propensity score.

Table 3.16. Comparability of Groups Formed by Pair Matching. Categorical Variables.

Variable	Level	Chemo	No Chemo	LR χ^2
Nodal	Negative	168	163	0.6815
	Positive	132	137	
ER	Pos	193	193	0.7888
	Neg	88	84	
	Unk	19	23	
Grade	1	40	37	0.9198
	2	180	184	
	3	80	79	
Year	1985	20	19	0.8807
	1988	26	32	
	1992	48	53	
	1995	70	66	
	1998	63	54	
	2001	73	76	
T4	Yes	15	17	0.7162
	No	285	283	

Table 3.17. Comparability of Groups Formed by Pair Matching. Continuous Variables.

Variable	Chemo	Mean	SD	p-value
Ln(size)	Yes	3.1915	0.5788	0.8354
	No	3.2015	0.5978	
Age	Yes	57.320	11.071	0.7947
	No	57.577	12.996	

It is interesting to note that the patients in the matched pairs are those with intermediate propensity scores, and as such form a population that would be clinically considered to be at intermediate risk of recurrence. A little more than half of them are node-negative, and almost two thirds are estrogen-receptor positive. Average tumour size is also intermediate, at about 25 mm.

3.6.3 Survival Analysis

Figure 3.5 shows the Kaplan-Meier curve comparing the overall survival of these two groups. No adjustments were made to the log rank test because there were no imbalanced variables between the groups.

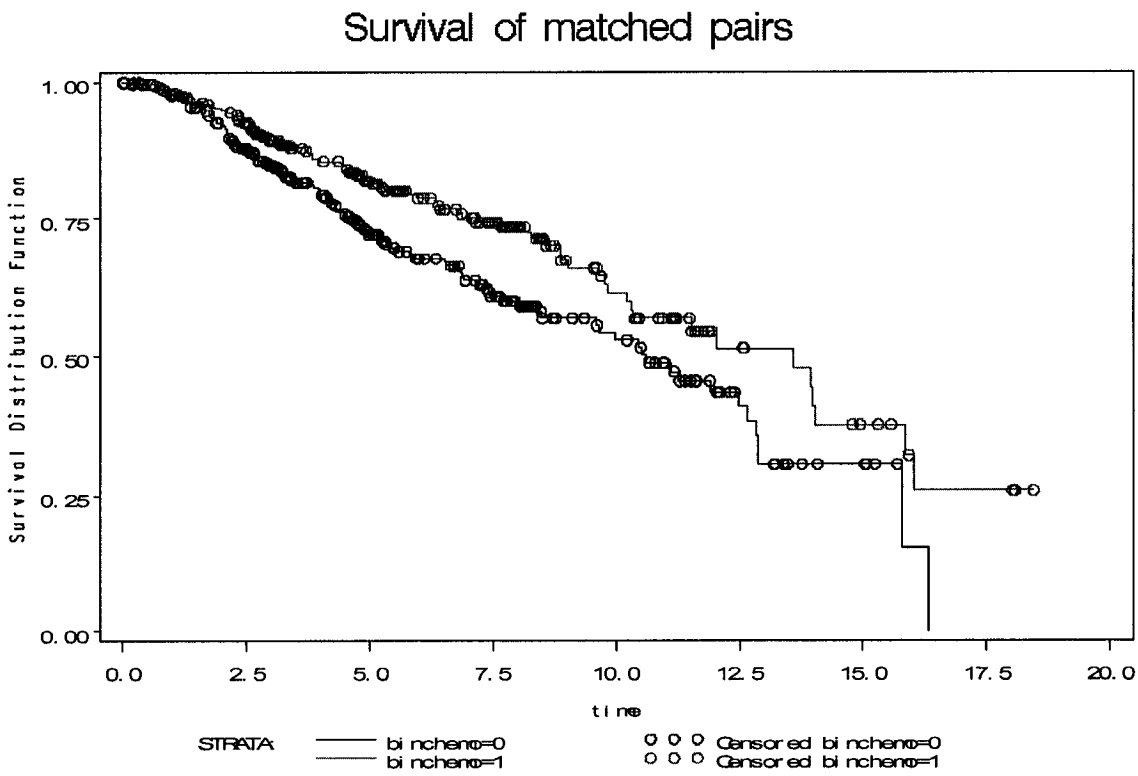


Figure 3.5. Overall survival of groups defined by pair matching process. Log rank p=0.0060. HR=0.662 (95% CI 0.439-0.891).

Paired survival analysis was undertaken by conditional logistic regression, using a unique number assigned to each pair as the stratification variable. The results of this paired analysis are only minimally different from the unpaired analysis (HR=0.646 95% CI 0.454-0.918).

3.7 Conventional Survival Analysis

To further characterize the value of the propensity score a survival analysis was undertaken without the use of the propensity score. This analysis was performed following the method described in section 2.7.5.

Univariate survival analysis is summarized in table 3.17. The variables significantly associated with survival are variables comprising the TNM stage, histologic grade, and hormone receptor status. Interestingly, in the univariate setting, neither receipt of chemotherapy, receipt of hormone therapy, nor year of treatment is associated with overall survival.

Continuous variables included in the model were age at diagnosis, number of nodes positive, and size of primary tumour. Details of the assessment of the form of these variables can be found in the supplementary tables (3.11s-3.13s) and graphs (graph 3.9s-3.11s). Tumour size was usable without transformation, number of positive nodes was optimized by taking the square root of all values, while age at diagnosis could not be transformed into a format that was superior to dividing it into quintiles. The interaction terms that are included are those for interactions between hormone therapy and histologic grade, and between number of positive nodes and size of tumour.

The results of the multivariable regression analysis are found in table 3.18. After other factors are adjusted for, chemotherapy in this patient population was associated with a 21.7% decreased hazard of death from any cause (Table 3.19).

It is interesting to note that the year of treatment was not significantly associated with overall survival. It is also interesting to note that the number of positive nodes could be modeled as a continuous variable when determining its effect on overall survival, but had to be dichotomized (into node-positive and node-negative) for purposes of modeling its effect on whether or not chemotherapy was recommended. Finally, it is not particularly surprising that there is an interaction between hormone therapy and the histologic grade of the tumour, but it is surprising that this interaction is more significant than the interaction between hormone therapy and estrogen receptor status, which did not make it into the final model.

Table 3.18. Results of Univariate Survival Analysis: Effect of baseline variables on overall survival.					
Variable	Level	Median Survival	logrank p	HR	95% CL
Nodal	Pos	10.32	<0.0001	1.615	1.412-1.848
	Neg	13.64			
Side	Left	12.97	0.5090	0.956	0.863-1.093
	Right	12.48			
Hormone	No	13.1718	0.1197	1.116	0.972-1.283
	Yes	10.84			
Chemo	No	12.65	0.8903	1.011	0.871-1.173
	Yes	13.26			
T4	No	12.90	<0.0001	3.046	2.424-3.828
	Yes	4.88			
Grade	1	13.19	<0.0001		
	2	12.65		1.670	1.341-2.080
	3	11.70		2.203	1.713-2.834
ER	Neg	13.98	<0.0001		
	Pos	12.64		0.685	0.582-0.807
	Unk	12.80		0.802	0.647-0.995
Prog	Neg	12.67	<0.0001		
	Pos	12.86		0.680	0.583-0.792
	Unk	12.76		0.854	0.698-1.045
Path Stage	1	14.17	<0.0001		
	2A	12.90		1.393	1.174-1.653
	2B	9.80		2.047	1.695-2.473
	3A	8.27		2.753	2.089-3.628
	3B	4.89		4.072	3.174-5.224
T-stage	1A	14.75	<0.0001		
	1B	14.07		1.753	0.972-3.160
	1C	13.01		2.460	1.410-4.293
	2	11.60		3.499	2.010-6.089
	3	8.90		5.021	2.786-9.048
	4A	3.86		10.826	5.255-22.303
	4B	6.63		7.085	3.695-13.583
	4C	3.40		10.488	4.784-22.993
	4D	4.88		8.466	4.287-16.719
N-stage	0	14.067	<0.0001		
	1a	9.00		1.182	0.631-2.214
	1b1	12.68		1.482	1.236-1.778
	1b2	8.96		2.367	1.846-3.035
	1b3	8.41		2.115	1.611-2.779
	1b4	9.58		2.015	1.397-2.891
	2	4.52		3.499	2.297-5.330
	3	2.98		2.776	0.691-11.146
	Unknown	6.38		2.957	2.345-3.370
Year	1985	10.31	0.3655		
	1988	13.30		0.848	0.678-1.059
	1992	11.99		0.884	0.704-1.109
	1995	NR		1.002	0.786-1.277
	1998	NR		0.877	0.664-1.158
	2001	NR		0.760	0.535-1.078

Table 3.19. Results of Cox Proportional Hazards Regression: Effect of baseline variables on overall survival					
Variable	**Level**	**Parameter Estimate**	**Standard Error**	**Hazard Ratio (95% CI)**	**p-value**
Chemotherapy	Yes vs. No	-0.24412	0.10558	0.783 (0.637-0.963)	0.0208
Hormone Therapy	Yes vs. No	-0.51930	0.21642	0.595 (0.389-0.909)	0.0164
Age	Rank 2 vs. 1	-0.10757	0.12575	0.898 (0.702-1.149)	0.3923
	Rank 3 vs. 1	0.18011	0.12599	1.197 (0.935-1.533)	0.1528
	Rank 4 vs. 1	0.49902	0.12432	1.647 (1.291-2.102)	<0.0001
	Rank 5 vs. 1	1.18097	0.12399	3.245 (2.555-4.154)	<0.0001
Nodes positive	Continuous	0.36746	0.05541	1.444 (1.295-1.610)	<0.0001
Size of tumour (mm)	Continuous	0.01369	0.00200	1.014 (1.010-1.018)	<0.0001
T4	Yes vs. No	0.45401	0.13220	1.575 (1.215-2.040)	0.0006
Estrogen Receptors	Pos vs. Neg	-0.39612	0.09520	0.673 (0.558-0.811)	<0.0001
	Unk vs. Neg	-0.26129	0.11898	0.770 (0.610-0.972)	0.0281
Grade	Int vs. Low	0.19251	0.14527	1.212 (0.912-1.612)	0.1851
	High vs. Low	0.61112	0.17783	1.842 (1.300-2.611)	0.0006
Hormone*grade	1 vs. 0*	0.48186	0.22906	1.619 (1.033-2.537)	0.0304
	2 vs. 0*	0.12951	0.26636	1.138 (0.675-1.919)	0.6268
Nodes*size	Continuous	-0.00178	0.0008641	0.998 (0.997-1.000)	0.0391

*Hormone therapy was coded as either 0 or 1 (not given or given). Grade was coded as 0, 1, or 2 (Low, intermediate, or high) so the variable created by their product has a value of 0 if the patient *either* has a low grade tumour, *or* got no hormonal therapy, a value of one if they had an intermediate grade tumour *and* hormonal therapy, and a value of two if they had a high grade tumour *and* hormonal therapy.

4.0 Discussion

This thesis was undertaken with the intent of determining the relative contribution of mammographic screening and adjuvant chemotherapy to the improvement in survival seen in breast cancer patients over the last twenty years. A secondary intent of the thesis was to apply propensity score analysis to this data, and to understand the merits and disadvantages of this approach as compared to conventional proportional hazards modeling.

4.1 Consistency of Results with Previously Published Data

The results of this review are quite consistent with the results of the EBCTCG meta-analysis of studies of adjuvant chemotherapy. The majority of studies included in this meta-analysis were ongoing during the same years analyzed in this thesis. See table 4.1 for a comparison of the survival times in this study with those in the 1998 meta-analysis. Although the meta-analysis has been updated since 1998 to include 15 year follow up, those updates report breast cancer specific mortality, rather than all-cause mortality, and hence cannot be directly compared to this thesis.

For most groups the point estimate from the meta-analysis is within the confidence intervals of this thesis. The two groups that have divergent survival results are younger node-positive women treated with chemotherapy and node-negative women not treated with chemotherapy, regardless of age.

In the case of younger node-positive women treated with chemotherapy, they had better survival in my series than in the meta-analysis. This may be because many of the

Table 4.1. Comparison of OHRCC* and EBCTCG* Results						
Age	Nodal Status	Chemo	5-Year OS (95% CL)		10-Year OS (95% CL)	
			EBCTCG**	OHRCC	EBCTCG**	OHRCC
<50	Negative	Yes	86.5%	85.7% (79.4-92.0)	77.6%	72.4% (60.4-84.5)
		No	83.5%	89.7% (85.7-93.8)	71.9%	79.1% (72.6-85.5)
<50	Positive	Yes	68.8%	78.0% (73.2-83.3)	53.8%	62.9% (55.9-69.8)
		No	61.8%	62.3% (32.9-91.7)	41.4%	33.3% (0-67.4)
50-69	Negative	Yes	85.3%	84.6% (77.6-91.7)	71.2%	73.3% (60.7-85.8)
		No	81.4%	90.7% (88.4-92.9)	64.8%	74.4% (71.8-78.8)
50-69	Positive	Yes	70.8%	75.1% (68.8-81.4)	48.6%	50.2% (38.9-61.4)
		No	61.7%	72.7% (65.9-79.4)	46.3%	52.4% (44.0-60.8)

* OHRCC – Ottawa Hospital Regional Cancer Centre EBCTCG – Early Breast Cancer Trialists Cooperative Group
**From Reference 74. Precise confidence limits cannot be ascertained from the data presented in the paper.

trials included in the meta-analysis required women to have at least four positive lymph nodes, while my series includes all node-positive women, 66% of whom had three or fewer nodes positive. Therefore, my node-positive population may have had less aggressive disease than the node-positive population in the meta-analysis.

My data includes many women whose risk of recurrence is thought to be so low that they would never have been enrolled in clinical trials. These women are particularly included in the group with node-negative disease. Trials of chemotherapy in node-negative women tend to enroll patients with high-risk features, such as large tumour size, high histologic grade, or absence of estrogen receptors. In my population, only 25% of node-negative patients were ER negative, only 18% were high grade, and 86% had tumours of three centimeters or smaller. The inclusion of these patients likely accounts for the better survival in the OHRCC population than in the EBCTCG population. The

inclusion of these patients may also account for the lack of a survival advantage associated with adjuvant chemotherapy in node-negative patients in this thesis. The lack of an advantage may not relate to treated patients doing more poorly than those in trials, but to untreated patients doing better than untreated trial patients.

Table 4.1 demonstrates that even when the results between the OHRCC population and the EBCTCG population are similar, the OHRCC one is usually slightly better. This may reflect the slight bias within OPIS to default to coding patients as "Alive" (See section 2.2).

4.2 The Effect of Mammography

Information about mammography use for individual patients is not available in OPIS. Although data linking mammography to individuals exists in provincial databases of physician billing, this data likely would not have helped to address the issue of screening mammography because physician billing codes make no distinction between mammography performed for screening and mammography performed for diagnosis of a palpable breast mass.

Therefore, the effect of mammography was inferred from the decrease in average tumour size over the interval that screening mammography went from an experimental intervention to a near-ubiquitous procedure. In our database the mean tumour size decreased from 28.12 mm to 21.86 mm between 1985 and 1998. In 2001 the mean size increased to 26.02 mm. This increase in size in the last year of the study may be an anomaly. An alternate hypothesis is that as it became more widely used, screening mammography identified a population of women with indolent, otherwise subclinical breast cancers, and that the pool of women with these cancers was depleted by 2001. An

analogous phenomenon was observed after the introduction of PSA screening for prostate cancer. Arguing against this hypothesis, however, is the fact that this hypothetical population of women would contribute to an increase in the absolute number of new diagnoses seen, and the number of new cases seen annually at the OHRCC is fairly constant, increasing by only 5.4% over the interval 1992-2001 (Table 3.1).

Although the cause of the decrease in tumour size over time cannot be ascertained from the present data, the decrease in mean tumour size corresponds with the increasing adoption of screening mammography in the general population (see section 1.2). It is possible that other factors may contribute to this phenomenon. Over the time period of this study women and physicians may have become more aware of the importance of breast exams and therefore detect earlier, smaller tumours. It is also possible that some unknown environmental factor involved in the pathogenesis of breast cancer has been changed or attenuated in recent years, resulting in tumours with slower growth kinetics that are more likely to be small when detected at a screening investigation.

4.3 The Effect of Adjuvant Chemotherapy

Over the interval covered in this thesis, the fraction of patients receiving no adjuvant systemic therapy decreased from 74% to 18%. The fraction receiving chemotherapy increased from 17% to 48% between 1985 and 2001, and of those receiving chemotherapy, the fraction receiving anthracyclines increased from 7% to 98%.

By 1992 almost all patients who had tumours expressing estrogen receptors were being treated with adjuvant hormonal therapy, so it is difficult to separate these two factors in the propensity score modeling. However, it was possible to use both of these variables in the Cox proportional hazards model to give some idea of the advantage of

hormonal therapy distinct from the survival advantage conveyed by having a biologically less aggressive tumour.

The stratified propensity score analysis demonstrates that a subset of the population in this study derives an improvement in overall survival from adjuvant chemotherapy. This group is in the top three cohorts of the propensity-stratified node-positive population. The characteristics of this group make it unsurprising that they derive the most benefit: they are young (median age 48.54 years) and therefore have few competing causes of mortality. They also have fairly aggressive cancers, with 26% of them having more than 5 positive lymph nodes.

The stratified propensity score analysis did not demonstrate any population of node-negative patients who derived an overall survival benefit from chemotherapy, though there was a strong trend towards benefit in the highest-risk strata (HR 0.661 95% CI 0.399-1.096).

The matched propensity score analysis also demonstrates an overall survival advantage associated with chemotherapy. This is interesting because the patients included in this analysis had generally less aggressive tumours than those in the higher strata of the node-positive stratified analysis. Almost half (45%) of these patients were node-negative, with an average tumour size of 24 mm and an average age of 57 years. See below for a discussion of why different patient populations are represented in the stratified and matched analyses.

Interestingly, the propensity score analysis also suggests that there is a population of women who may be harmed by adjuvant chemotherapy. In the first cohort of the node-positive population chemotherapy administration was associated with an adjusted

hazard for death of 2.313 (95% CI 1.236-4.326). The women in this cohort had a mean age of 73.9 years, and 46% of them had only one node positive. In addition, 85% had tumours expressing estrogen receptors. In routine clinical practice these patients would be considered borderline candidates for systemic chemotherapy, and this small retrospective review gives further cause to think carefully before suggesting it to these women.

4.4 Changes in Survival Over Time

This thesis was conceived to explore the reasons why breast cancer mortality has been declining in recent decades, with particular focus on mammographic screening and the increase in adjuvant chemotherapy. The data demonstrate a decrease in tumour size over time, and a survival advantage associated with adjuvant chemotherapy. Therefore, it is surprising that there is no apparent association between year of treatment and survival, either in univariate analysis or in the Cox proportional hazards model.

Although this thesis includes analysis of many patients with long follow up time, the mean follow up time is only 5.88 years. This is because the earlier cohorts have relatively few patients, and the more recent cohorts have many patients but short follow up. For instance, 22% of patients were in the 2001 cohort, a cohort that had no more than 3.5 years of follow up at the time the database was constructed. Therefore, the ability to detect a difference in overall survival over time is hampered by small numbers of events in the earlier cohorts, and short follow-up in later cohorts.

In the Cox proportional hazards model, each decrease in tumour size of one millimeter was associated with a decrease in the hazard of death of 1.4%. Thus, the decrease in mean tumour size between 1985 and 1998 should have been associated with

8.8% decrease in the hazard of death, or a 2.9% decrease in the absolute risk of death at ten years in the node-positive group. Detection of an effect of this size would have required many more patients than this study describes.

Given the dramatic increase in the use of adjuvant chemotherapy over the time period described in this study, and given also that chemotherapy is shown to be associated with a survival advantage in appropriately selected patients, it is puzzling that no relationship was seen between year of treatment and survival. This may be because the three most recent cohorts do not yet have long enough follow up to demonstrate improved survival when compared to earlier cohorts. A second possibility is that as the indications for adjuvant chemotherapy have expanded into lower-risk patients the additional patients derive a smaller benefit from their treatment. If this is the case then the increased use of chemotherapy over time will not necessarily be associated with an increase in survival of equal magnitude.

The question that is central to this discussion is whether or not this study was sufficiently powered to detect the changes in survival over time that might be associated with increasing use of adjuvant chemotherapy. We have already seen that the improvement in overall survival associated with decreasing tumour size over time is likely too small to be demonstrated by this study. The same is likely true of the effect of adjuvant chemotherapy.

One can estimate the power to detect a difference between the earliest (1985) and latest (2001) cohort in this study by using the sample size calculations conventionally done for a clinical trial. Using the assumptions of a 19.5 year maximum follow up, a median untreated survival of 12.64 years and a hazard ratio of 0.783 (from table 3.18)

between the two arms, we find that 741 patients would be required to demonstrate this difference with α=0.05 and β=0.2, slightly less than the actual number of 887 in those two cohorts in this study. However, this calculation is applicable to a clinical trial, and the present retrospective analysis has lower power than a clinical trial for several reasons: the first is that although the maximum follow up is 19.5 years, the follow up is different in each arm, being up to 19.5 years in the 1982 cohort, but no more than 3.5 years in the 2001 cohort. As well, the conventional sample size calculation assumes that patients are randomized 1:1 to treatment or observation. In this case the patients are distributed 2.8:1 to the 2001 cohort. Finally, the magnitude of the hazard ratio between the treatment-averse 1985 cohort and the highly-treated 2001 cohort is unlikely to be as great as 0.783, the difference between completely treated and untreated populations. All of these factors erode statistical power and suggest that this study was not sufficiently powered to demonstrate changes in survival related to differences in chemotherapy use between the earliest and latest cohorts.

This study did demonstrate a decrease in tumour size accompanied by an increase in the use of adjuvant chemotherapy over the study period. It also demonstrated improvements in overall survival associated with smaller tumour size and with adjuvant chemotherapy. It was, however, likely underpowered to demonstrate the expected change in overall survival related to these two factors between the beginning and the end of the study period.

4.5 Propensity Score Analysis versus Standard Proportional Hazards Modeling

The propensity score attempts to account for several baseline covariates in a single number in order to facilitate stratification or matching on multiple variables.

Multivariable Cox proportional hazards modeling accounts for all variables in the model simultaneously through calculations significantly more complex than those involved in the use (though not the generation) of propensity scores. For the purposes of this thesis, two important limitations of propensity scores were identified.

The first problem with propensity scores is that they are not very well suited to illustrating the effect of changes in therapy over time. This question was easily addressed in the Cox model by including interaction terms between year of diagnosis and tumour size, or year of diagnosis and type of treatment. Similar analysis with propensity scores would have required a separate propensity model for each year covered in the study and would likely run into trouble with small patient numbers once the patients in each year were further divided into quintiles on the basis of their propensity scores.

The second significant problem with propensity scores, particular to this study, was the bimodal distribution of scores in the study population. Fifty five percent of patients had propensity scores either less than 0.025 or greater than 0.975. This makes stratification on propensity scores dubious, and makes it very likely that the strata with the lowest and highest scores will have too few patients in either the treatment or non-treatment arms to allow meaningful comparison. It also means that a matched case-control analysis will exclude most patients from the extremes of the propensity score spectrum and include mainly patients with intermediate propensity scores. This subset of patients may thereby be inappropriately emphasized.

All this being said, however, the results of the propensity score analysis do give information that is different from the results of the Cox modeling. The Cox model tends to group all patients together and implies a continuous gradient of benefit for all members

of the group, with the magnitude of benefit (or risk) based on an individual's particular baseline covariates. In contrast, the stratified propensity score analysis divides the population into groups and determines the benefit of an intervention within each group. This may be a more reasonable approach, particularly if the groups are clinically distinct. In this case the stratified propensity score analysis gives more information about the effect of chemotherapy on node-negative versus node-positive patients than does the Cox model.

It is more difficult to compare the results of the matched propensity score analysis to the conventional Cox modeling because the matched analysis looked at a population of intermediate-risk patients whose inclusion was an artifact of the bimodal distribution of propensity scores, rather than being dictated by some relevant clinical variables. The results of the matched analysis are consistent with the other analysis, though, as the hazard ratio associated with chemotherapy in this group (0.662) fell between the hazard ratios for the highest-risk node-positive patients (0.392) and the lowest-risk node-negative patients (1.00), and was slightly lower than the hazard ratio associated with chemotherapy for the entire population (0.783) in the Cox model.

Thus, all three modeling strategies provide coherent and consistent results. The difference in the observed effect of adjuvant chemotherapy in each group can be attributed to differences in the baseline risk of the population included in each analysis. The addition of the propensity score analysis did not alter the conclusions of the thesis, but was more effective than the Cox analysis at illustrating the benefit of adjuvant therapy within different patient subgroups.

It is likely that some amount of the observed benefit to chemotherapy administration represents selection bias. Many retrospective reviews, including this one have no means of quantifying patient co-morbidity. It is likely that some of the untreated patients with high propensity scores were untreated because of other severe medical conditions that limited their lifespan. It is not possible with the current data to separate the effects of selection bias from the effects of chemotherapy administration on survival. The effect of selection bias is probably somewhat mitigated, however, by a tendency for referring physicians not to send patient to the OHRCC for consideration of adjuvant therapy if they are profoundly medically unwell for other reasons.

4.6 Significance of Results

This is the largest population-based study of outcomes in early breast cancer patients to incorporate individual patient chart review. Any larger studies have been based on SEER data, BCCA registry data, or other large administrative databases that contain less detail about staging and treatment. (see section 1.4)

In general, the value of population-based analysis lies in the usefulness of demonstrating that treatment effects seen in clinical trials are also observable in the general population. Clinical trials enroll selected patients and tend to include measures to ensure treatment compliance and follow-up that do not exist in standard practice. For this reason there are concerns that the benefits of a therapy in a clinical trial may be attenuated when that therapy is applied to the general population. This concern is particularly relevant for cytotoxic chemotherapy, an intervention that can have a narrow therapeutic index.

This thesis demonstrates that general patients with node-positive cancers derive similar benefit from adjuvant chemotherapy as do clinical trials patients, as reported in the EBCTGC meta-analysis. Given this finding, it would be surprising if node-negative patients did not also derive benefit from adjuvant chemotherapy, as they do in clinical trials. This benefit is smaller, however, than the benefit for node-positive patients, and the present thesis may be underpowered to demonstrate it. It is interesting to note that treated node-negative patients had comparable outcomes in the meta-analysis and in this thesis. Untreated patients did better in this thesis than they did in the meta-analysis, suggesting that the unselected nature of this population includes many women with very low-risk disease who were not treated. Because the non-treated group does so well, the hazard ratio associated with chemotherapy administration tends towards one.

4.7 Summary

This thesis demonstrates that the benefits of adjuvant chemotherapy seen in clinical trials are also demonstrable in a population of unselected patients seen at a tertiary cancer centre. It was not possible to demonstrate a relationship between year of treatment and improved outcomes, as the study population was probably underpowered to demonstrate such a relationship.

5.0 Glossary of Terms

5.1 Glossary of Statistical Terms

c-statistic: The c statistic is a measure of the discriminative power of the logistic equation. It varies from .5 (the model's predictions are no better than chance) to 1.0 (the model always assigns higher probabilities to correct cases than to incorrect cases for any pair involving dependent=0 and dependent=1). Thus c is the percent of all possible pairs of cases in which the model assigns a higher probability to a correct case than to an incorrect case.

Conditional logistic regression: A modification of standard logistic regression which is better suited to highly stratified data. It is commonly used for matched pair designs where each pair represents a stratum.

Hosmer and Lemeshow goodness-of-fit test: A test of model fitting in logistic regression. The test divides subjects into deciles based on predicted probabilities, then computes a chi-square from observed and expected frequencies. Then a p-value is computed from the chi-square distribution with 8 degrees of freedom to test the fit of the logistic model. Well-fitting models show nonsignificance on this test.

Loess: A type of moving-average non-parametric regression. In each interval a least-squares linear regression is performed, with increased weight given to the values at the centre of the interval.

Martingale Residuals: A residual in logistic regression that has an expectation of zero, and can be thought of by the formulation 'observed' minus 'expected'.

Propensity score: The predicted probability of being in the treatment (versus control) group, a propensity score is determined by incorporation of baseline covariates into a multivariable logistic regression model. The score reduces each patient's set of covariates to a single number, making it possible to match or stratify on multiple variables simultaneously.

Quasi-complete separation of data: In logistic regression, this situation occurs when one level of a categorical variable includes very few individuals. In this situation the maximum likelihood estimate is very imprecise, or may not exist.

5.2 Glossary of Clinical Terms

Adjuvant therapy: Therapy administered after a potentially curative surgical resection of a cancer. Patients are free of observable disease, and the intent of adjuvant therapy is to eliminate microscopic residual cancer and increase the likelihood of cure.

Early breast cancer: Breast cancer restricted to the breast and also possibly to the lymph nodes in the underarm (axilla) on the same side as the affected breast. Early breast cancer is treated with curative intent. Treatment includes surgery, often followed by adjuvant radiation, chemotherapy, or hormonal therapy.

Grade: Also termed *histologic grade*. A measure of how abnormal cancer cells look under the pathologist's microscope. Cells that look similar to non-cancer cells are termed low-grade, while those with a more abnormal appearance are called high-grade. Commonly used grading systems in breast cancer include the Fisher Nuclear Grade, and the Scarff-Bloom-Richardson Score.

Metastatic breast cancer: Cancer that has spread beyond the breast and its regional lymph nodes, often into the bones, lungs, liver or brain. Metastatic breast cancer is generally incurable.

Node-negative: Description of a tumour that has not spread from the breast into regional lymph nodes.

Node-positive: In the case of breast cancer, a tumour whose cells have spread away from the primary tumour and into lymph nodes, usually under the arm on the same side as the affected breast. Spread into other, more distant lymph nodes is considered metastatic breast cancer.

Screening: Testing for cancer in an asymptomatic population in the hopes of detecting cancer at an early, more treatable stage. Commonly used cancer screening tests include mammography for breast cancer, Pap smears for cervical cancer, PSA testing and digital rectal exam for prostate cancer, and colonoscopy or barium enema for colon cancers.

Therapeutic Index: For a medication, the ratio of the dose required to help a certain percentage of patients to the dose required to harm a certain percentage of patients. Safer medications have a large (often called *wide*) therapeutic index, less safe medications a small (or *narrow*) therapeutic index.

6.0 References

1. National Cancer Institute of Canada: Canadian Cancer Statistics 2002, Toronto, Canada, 2002; 26-32.
2. Reynolds, T. Declining Breast Cancer Mortality: What's Behind It? JNCI 1999; 91:750-3.
3. Jatoi I, and Miller, AB. Why is Breast Cancer Mortality Declining? Lancet Oncol 2003; 4:251-4.
4. DeGrasse CE, O'Connor AM, Perrault DJ, et al. Changes in women's breast cancer screening practices, knowledge, and attitudes in Ottawa-Carlton since 1990. Can J Public Health. 1996; 87(5):333-8.
5. DeGrasse CE, O'Connor AM, Boulet J, et al. Changes in Canadian women's mammography rates since the implementation of mass screening programs. Am J Public Health. 1999; 89(6):927-9.
6. Gaudet LA, Altmayer CA, Nobrega KM et al. Trends in mammography utilization, 1981 to 1994. Health Rep. 1996 8(3):17-27 (Eng) 19-30 (Fre)
7. The Workshop Group. Reducing deaths from breast cancer in Canada. Can Med Assoc J. 1989; 141:199-201.
8. Health Canada: Organized Breast Cancer Screening Programs in Canada 1999 and 2000 Report. Toronto, Canada, 2001; 13.
9. Tabar L, Fagerberg G, Duffy S et al. The Swedish two-county trial of mammographic screening for breast cancer: Recent results and calculation of benefit. J Epidemiol Community Health. 1989; 43:107.
10. Andersson I, Aspegren K, Janzon L et al. Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial. BMJ. 1988; 297:943.
11. Alexander FE, Anderson TJ Brown HK et al. The Edinburgh randomized trial of breast cancer screening: results after 10 years of follow up. Br. J Cancer. 1994; 70: 542.
12. Chu KC, Smart CR, Tarone RE, et al. Analysis of breast cancer mortality and stage distribution by age for the Health Insurance Plan clinical trial. J Natl Cancer Inst. 1988 Sep 21;80(14):1125-32.
13. Frisell J, Eklund G, Hellstrom L, et al. A. Randomized study of mammography screening—preliminary report on mortality in the Stockholm trial. Breast Cancer Res Treat. 1991 Mar;18(1):49-56.
14. Bjurstam N, Bjorneld L, Warwick J, et al. The Gothenburg Breast Screening Trial. Cancer. 2003 May 15;97(10):2387-96.
15. Miller AB, Baines CJ, To T, et al. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. CMAJ. 1992 Nov 15;147(10):1459-76.
16. Miller AB, Baines CJ, To T, et al. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. CMAJ. 1992 Nov 15;147(10):1477-88.
17. Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL. Efficacy of screening mammography. A meta-analysis. JAMA. 1995 Jan 11;273(2):149-54
18. Hackshaw A. EUSOMA review of mammography screening. Ann Oncol. 2003 Aug;14(8):1193-5.

19. Nystrom L, Rutkvist LE, Walls S, et al. Breast cancer screening with mammography: overview of Swedish randomized trials. Lancet. 1993; 341:973-8.
20. Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? Lancet. 2000 Jan 8;355(9198):129-34.
21. Alexander F, Roberts MM, Lutz W, et al. Randomisation by cluster and the problem of social class bias. J Epidemiol Community Health. 1989; 43:29-36.
22. deKoning HJ. Mammographic screening: evidence from randomized controlled trials. Ann Oncol. 2003; 14:1185-9.
23. Hoel DG, Ron E, Carter R, et al. Influence of death certificate errors on cancer mortality trends. J Natl Cancer Inst. 1993 Jul 7;85(13):1063-8.
24. Prorok PC. Epidemiologic approach for cancer screening. Problems in design and analysis of trials. Am J Pediatr Hematol Oncol. 1992 May;14(2):117-28.
25. Olivotto IA, Mates D, Kan L, et al. Prognosis, treatment, and recurrence of breast cancer for women attending or not attending the Screening Mammography Program of British Columbia. Breast Cancer Res Treat. 1999 Mar;54(1):73-81.
26. Eckhardt S, Badellino F, Murphy GP. UICC meeting on breast cancer screening in premenopausal women in developed countries. Int J Cancer. 1994; 56:1.
27. Elwood JM, Cox B, Richardson AK. The effectiveness of breast cancer screening by mammography in younger women. Online J Curr Clin Trials. 1993; Doc#32.
28. Smart CR, Hendrick RE, Rutledge JH, et al. Benefit of mammography screening in women aged 40-49 years: current evidence from randomized controlled trials. Cancer 1995; 75:1619.
29. Hendrick RE, Smith RA, Rutledge JH, et al. Benefit of screening mammography in women aged 40-49: a new meta-analysis of randomized controlled trials. J Natl Cancer Inst Monogr. 1997; 22:87.
30. Berry DA. Benefits and risks of screening mammography for women in their forties: a statistical appraisal. J Natl Cancer Inst. 1998; 90:1431.
31. Beatson GT. On the treatment of inoperable cases of carcinoma of the mamma: suggestions for a new method of treatment with illustrative cases. Lancet. 1896; 2:104-7.
32. DeVita VT, Hemmlan S, Rosenberg SA, eds. Cancer: Principles and Practice of Oncology. 6th ed. Lippincott Williams & Wilkins, 2001. p1635
33. Kaufmann M. Zoladex (goserelin) vs. CMF as adjuvant therapy in pre/perimenopausal node-positive early breast cancer. Preliminary efficacy results from the ZEBRA study. Breast. 2001; 10(Suppl 1): s30.
34. Wallweiner D, Possinger K, Bondar G, et al. Leuproelin acetate vs. CMF in the adjuvant treatment of premenopausal women with ER/PR-positive node-positive breast cancer: interim results of the TABLE study. Proc ASCO. 2001; 20:34a (abstr 132)
35. Jakesz R, Hausmaninger H, Samonigg H, et al. Complete endocrine blockade with tamoxifen and goserelin is superior to CMF in the adjuvant treatment of premenopausal lymph node-positive and -negative patients with hormone-responsive breast cancer. Breast. 2001; 10(Suppl 1): s10.
36. Boccardo F, Rubagotti A, Amoroso D, et al. Cyclophosphamide, methotrexate, fluorouracil versus tamoxifen plus ovarian suppression as adjuvant treatment of estrogen receptor-positive pre-/perimenopausal breast cancer patients: results of the Italian Breast Cancer Adjuvant Study Group 02 Randomized Trial. J Clin Oncol. 2000; 18: 2718-2727.

37. Roche HH, Kerbrat P, Bonneterre J, et al. Complete hormonal blockade versus chemotherapy in premenopausal early-stage breast cancer patients with positive hormone-receptor and 1-3 node positive tumour: results of the FASG 06 trial. Proc ASCO. 2000; 19: 72a (abstr 279).
38. Baum M, Houghton J, Odling-Smee W, et al. Adjuvant Zoladex in premenopausal patients with early breast cancer: results from the ZIPP trial. Breast. 2001; 10(Suppl 1): s32-33.
39. Davidson N, O'Neill A, Vukov CK, et al. Effectg of chemo-hormonal therapy in premenopausal node-positive, receptor-positive breast cancer: an Eastern Cooperative Oncology Group phase III intergroup trial. Breast. 1999; 8: 232-233.
40. De Matteis A, Montedoro D, Nuzzo F, et al. Hormonotherapy with goserelin depot after adjuvant chemotherapy in pre-menopausal women with early breast cancer: Is there any benefit? Ann Oncol. 1998; 9 (Suppl 4): 11 (abstr 52).
41. International Breast Cancer Study Group. Adjuvant chemotherapy followed by goserelin versus either modality alone for premenopausal lymph node-negative breast cancer: A randomized trial. J Natl Cancer Inst. 2003; 95: 1833-46.
42. Pritchard KI. Adjuvant therapy for premenopausal women with breast cancer: Is it time for another paradigm shift? J Clin Oncol. 2002; 20:4611.
43. Baum M, Buzdar A, Cuzick J, et al. Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early-stage breast cancer: results of the ATAC (Arimidex, Tamoxifen Alone or in Combination) trial efficacy and safety update analyses. Cancer. 2003 Nov 1;98(9):1802-10.
44. Goss PE, Ingle JN, Martino S, et al. A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. N Engl J Med. 2003 Nov 6;349(19):1793-802.
45. Coombes RC, Hall E, Gibson LJ, et al. A randomized trial of exemestane after two to three years of tamoxifen therapy in postmenopausal women with primary breast cancer. N Engl J Med. 2004 Mar 11;350(11):1081-92.
46. Mouridsen H, Gershanovich M, Sun Y, et al. Superior efficacy of letrozole versus tamoxifen as first-line therapy for postmenopausal women with advanced breast cancer: results of a phase III study of the International Letrozole Breast Cancer Group. J Clin Oncol. 2001 May 15;19(10):2596-606.
47. Buzdar A, Douma J, Davidson N, et al. Phase III, multicenter, double-blind, randomized study of letrozole, an aromatase inhibitor, for advanced breast cancer versus megestrol acetate. J Clin Oncol. 2001 Jul 15;19(14):3357-66.
48. Nabholtz JM, Bonneterre J, Buzdar A, et al. Anastrozole (Arimidex) versus tamoxifen as first-line therapy for advanced breast cancer in postmenopausal women: survival analysis and updated safety results. Eur J Cancer. 2003 Aug;39(12):1684-9.
49. Kaufmann M, Bajetta E, Dirix LY, et al. Exemestane improves survival compared with megoestrol acetate in postmenopausal patients with advanced breast cancer who have failed on tamoxifen. results Of a double-blind randomised phase III trial. Eur J Cancer. 2000 Sep;36 Suppl 4:S86-7.
50. Early Breast Cancer Trialists Collaborative Group. Ovarian ablation in early breast cancer: overview of the randomized trials. Lancet. 1996; 348: 1189-96.

51. Riggs B. L., Hartmann L. C. Drug Therapy: Selective Estrogen-Receptor Modulators: Mechanisms of Action and Application to Clinical Practice. N Engl J Med 2003; 348:618-629.
52. Nolvadex Adjuvant Treatment Organisation. Controlled trial of tamoxifen as single adjuvant agent in early breast cancer: analysis at six years. Lancet. 1985; 1: 836-9
53. Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer (Cochrane Methodology Review). In: *The Cochrane Library,* Issue 4, 2003. Chichester, UK: John Wiley & Sons, Ltd.
54. Scottish Cancer Trials Office. Adjuvant tamoxifen in the management of operable breast cancer: The Scottish trial. Lancet. 1987; 2: 171-5.
55. Early Breast Cancer Trialist's Collaborative Group. Effects of adjuvant tamoxifen and of cytotoxic chemotherapy on mortality in early breast cancer. An overview of 61 randomized trials among 28,896 women. N Engl J Med. 1988; 319:1681-92.
56. Early Breast Cancer Trialists Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomized trials. Lancet. 1998; 351: 1451-67.
57. Members of the Breast Cancer Disease Site Group. Adjuvant Systemic Therapy for Node-negative Breast Cancer Practice Guideline Report #1-8. Cancer Care Ontario Practice Guidelines Initiative Web site. Available at: http://www.cancercare.on.ca/pdf/pebc1-8f.pdf. Accessed 9 June 2004.
58. The Steering Committee on Clinical Practice Guidelines for the Care and Treatment of Breast Cancer. Adjuvant systemic therapy for women with node-negative breast cancer. CMAJ. 1998 Feb 10;158 Suppl 3:S43-51.
59. The Steering Committee on Clinical Practice Guidelines for the Care and Treatment of Breast Cancer. Adjuvant systemic therapy for women with node-positive breast cancer. CMAJ. 1998 Feb 10;158 Suppl 3:S52-64.
60. Winer EP, Morrow M, Osborne CK, Harris JR. Malignant Tumours of the Breast. Principles and Practice of Oncology. Ed. Da Vita VT, Hellman S, Rosenberg SA. Philadelphia: Lippencott Williams and Wilkins. 2001. p 1692.
61. Early Breast Cancer Trialists' Collaborative Group. Multi-agent chemotherapy for early breast cancer (Cochrane Methodology Review). In: *The Cochrane Library,* Issue 4, 2003. Chichester, UK: John Wiley & Sons, Ltd.
62. Fisher B, Fisher ER, Redmond C. Ten-year results from the NSABP clinical trial evaluating the use of l-phenylalanine mustard (L-PAM) in the management of primary breast cancer. J Clin Oncol. 1986; 4:929-41.
63. Bonadonna G, Valagussa P, Rossi A, et al. Ten-year experience with adjuvant CMF-based chemotherapy in resectable breast cancer. Breast Cancer Res Treat. 1985; 5:95-115.
64. Tancini G, Bonadonna G, Valagussa P, et al. Adjuvant CMF in breast cancer: Comparative 5-year results of 12 versus 6 cycles. J Clin Oncol. 1983; 1:2-10.
65. Henderson IC, Gelman RS, Harris JR, et al. Duration of therapy in adjuvant chemotherapy trials. NCI Monogr. 1986; 1: 95-8.
66. Ahmann D, Bisel H, Eagan R, et al. Controlled evaluation of adriamycin (NSC-123127) in patients with disseminated breast cancer. Cancer Chemother Rep. 1974; 58:877-82.
67. Hoogstraten B, George SL, Samal B, et al. Combination chemotherapy and adriamycin in patients with advanced breast cancer. Cancer. 1976; 38:13-20.

68. Mathe G, Misset JL, Plagne R, et al. Adriamycin, vincristine, cyclophosphamide and 5-fluorouracil (AVCF) compared with cyclophosphamide, methotrexate and 5-fluorouracil (CMF) in premenopausal breast carcinoma. Personal results. Drugs Exp Clin Res. 1986;12(1-3):143-5.
69. Misset JL, di Palma M, Delgado M, et al. Adjuvant treatment of node-positive breast cancer with cyclophosphamide, doxorubicin, fluorouracil, and vincristine versus cyclophosphamide, methotrexate, and fluorouracil: final report after a 16-year median follow-up duration. J Clin Oncol. 1996 Apr;14(4):1136-45.
70. Fisher B, Redmond C, Wickerham DL, et al. Doxorubicin-containing regimens for the treatment of stage II breast cancer: The National Surgical Adjuvant Breast and Bowel Project experience. J Clin Oncol. 1989 May;7(5):572-82.
71. Martin M, Villar A, Sole-Calvo A, et al. Doxorubicin in combination with fluorouracil and cyclophosphamide (i.v. FAC regimen, day 1, 21) versus methotrexate in combination with fluorouracil and cyclophosphamide (i.v. CMF regimen, day 1, 21) as adjuvant chemotherapy for operable breast cancer: a study by the GEICAM group. Ann Oncol. 2003 Jun;14(6):833-42.
72. Coombes RC, Bliss JM, Wils J, et al. Adjuvant cyclophosphamide, methotrexate, and fluorouracil versus fluorouracil, epirubicin, and cyclophosphamide chemotherapy in premenopausal women with axillary node-positive operable breast cancer: results of a randomized trial. The International Collaborative Cancer Group. J Clin Oncol. 1996 Jan;14(1):35-45.
73. Levine MN, Bramwell VH, Pritchard KI, et al. Randomized trial of intensive cyclophosphamide, epirubicin, and fluorouracil chemotherapy compared with cyclophosphamide, methotrexate, and fluorouracil in premenopausal women with node-positive breast cancer. National Cancer Institute of Canada Clinical Trials Group. J Clin Oncol. 1998 Aug;16(8):2651-8.
74. Early Breast Cancer Trialists' Collaborative Group. Polychemotherapy for early breast cancer: an overview of the randomized trials. Lancet. 1998; 352: 930-42.
75. Crown J, O'Leary M, Ooi WS. Docetaxel and paclitaxel in the treatment of breast cancer: a review of clinical experience. Oncologist. 2004;9 Suppl 2:24-32.
76. Henderson IC, Berry DA, Demetri GD, et al. Improved outcomes from adding sequential Paclitaxel but not from escalating Doxorubicin dose in an adjuvant chemotherapy regimen for patients with node-positive primary breast cancer. J Clin Oncol. 2003 Mar 15;21(6):976-83.
77. Citron ML, Berry DA, Cirrincione C, et al. Randomized trial of dose-dense versus conventionally scheduled and sequential versus concurrent combination chemotherapy as postoperative adjuvant treatment of node-positive primary breast cancer: first report of Intergroup Trial C9741/Cancer and Leukemia Group B Trial 9741. J Clin Oncol. 2003 Apr 15;21(8):1431-9. Epub 2003 Feb 13.
78. Buzdar AU, Singletary SE, Valero V, et al. Evaluation of paclitaxel in adjuvant chemotherapy for patients with operable breast cancer: preliminary data of a prospective randomized trial. Clin Cancer Res. 2002 May;8(5):1073-9.
79. Pater JL, Pritchard KI. Optimal chemotherapy for women with breast cancer: the plot thickens. J Clin Oncol. 2003 Mar 15;21(6):963-4.
80. Mokbel K. Recent advances in breast cancer: the 25th Annual San Antonio Breast Cancer Symposium. Curr Med Res Opin. 2003;19(2):143-6.

81. Gaffney DK, Alexander T, Wiggins CL et al. Diminished survival in patients with inner versus outer quadrant breast cancers. J Clin Oncol 2003; 21(12):467-492.
82. Cooper GS, Virnig B, Klabunde CN et al. Use of SEER-Medicare data for measuring cancer surgery. Med Care 2002; 40(Suppl 8):IV43-IV48.
83. Virnig BA, Warren JL, Cooper GS et al. Studying radiation therapy using SEER-Medicare linked data. Med Care 2002; 40(Suppl 8):IV43-IV48.
84. Harlan LC, Hankey BF et al. The Surveillance, Epidemiology, and End-Results Program Database as a Resource for Conducting Descriptive Epidemiologic and Clinical Studies. J Clin Oncol 2003; 21(12):2232-2233.
85. Maggard MA, O'Connell JB, Lane KE, et al. Do young breast cancer patients have worse outcomes? J Surg Res 2003; 113(1):109-13.
86. Diab SG, Elledge RM, Clark GM. Tumour characteristics and clinical outcome of elderly women with breast cancer. J Natl Cancer Inst 2000; 92(7):550-6.
87. Shavers VL, Harlan LC, Stevens JL. Racial/ethnic variation in clinical presentation, treatment, and survival among breast cancer patients under age 35. Cancer. 2003; 97(1):134-47.
88. McCarthy EP, Burns RB, Coughlin SS et al. Mammography use helps to explain the differences in breast cancer stage at diagnosis between older black and white women. Ann Intern Med. 1998; 128(9): 729-36.
89. Li CI, Malone KE, Daling JR. Differences in breast cancer stage, treatment, and survival by race and ethnicity. Arch Internal Med 2003; 163(1):49-56.
90. Clegg LX, Li FP, Hankey BF. Cancer Survival among US whites and minorities: a SEER program population-based study. Arch Intern Med. 2002; 162(17):1985-93.
91. Henson DE, Chu KC, Levine PH. Histologic grade, stage, and survival in breast carcinoma: comparison of African American and caucasian women. Cancer 2003; 98(5): 908-17.
92. Chang S, Parker SL, Pham T, et al. Inflammatory breast carcinoma incidence and survival: the surveillance, epidemiology, and end results program of the National Cancer Institute, 1975-1992. Cancer 1998; 82(12): 2366-72.
93. Li CI, Moe RE, Daling JR. Risk of mortality by histologic type of breast cancer among women aged 50 to 79 years. Arch Intern Med. 2003; 163(18): 2149-53.
94. Chaudhry R, Goel V, Sawka C. Breast cancer survival by teaching status of the initial treating hospital. CNAJ 2001; 164(2): 183-8.
95. Paszat LF, Groome PA, Schulze K, et al. A population-based study of breast conservation for newly diagnosed breast cancer. Int J Radiat Oncol Biol Phys. 2000 Jan 15;46(2):345-53.
96. S. K. L. Chia, C. Speers, A. Kang, et al. The impact of new chemotherapeutic and hormonal agents on the survival of women with metastatic breast cancer (MBC) in a population based cohort. Proc Am Soc Clin Oncol 22: page 6, 2003 (abstr 22).
97. Hebert-Croteau N, Brisson J, Latreille J, et al. A population-based study of the impact of the St-Gallen treatment guidelines on survival of women with node-negative breast cancer. Breast. 2003 Mar;12 Suppl 1:S17-8.
98. Olivotto IA, Bajdik CD, Plenderleith IH, et al. Adjuvant systemic therapy and survival after breast cancer. N Engl J Med 1994; 330: 805-10.

99. Ugnat AM, Xie L, Morriss et al. Survival of women with breast cancer in Ottawa, Canada: variation with age, stage, histology, grade, and treatment. Br J Cancer. 2004; 90: 1138-43.
100. Ugnat AM. Personal communication.
101. Xie L, Ugnat AM, Morriss J, et al. Histology-related variation in the treatment and survival of patients with lung carcinoma in Canada. Lung Cancer. 2003; 42: 127-39.
102. Debevec L. Is 'no treatment' better than radiotherapy and chemotherapy? Lung Cancer. 2004; 45: 125.
103. Pisani P, Forman D. Declining mortality from breast cancer in Yorkshire, 1983-1998: extent and causes. Br J Cancer. 2004; 90: 652-6.
104. Chia KS, Du WB, Sankaranarayanan R, et al. Do younger female breast cancer patients have a poorer prognosis? Results from a population-based survival analysis. Int J. Cancer 2004; 108(761-95).
105. Jensen AR, Garne JP, Storm HH, et al. Stage and survival in breast cancer patients in screened and non-screened Danish and Swedish populations. Acta Oncol 2003; 42(7):701-9.
106. Brenner H, Hakulinen T. Are patients diagnosed with breast cancer before age 50 ever cured? J Clin Oncol 2004; 22(3): 432-8.
107. Weitzen S, Lapane KL, Toledano AY, et al. Principles for modeling propensity scores in medical research: A systemic literature review. Pharmacoepidemiology and Drug Safety 2004; 13:841-853.
108. . Shadish WR, Cook TD, Campbell DT. Experimental and Quasi-experimental Designs for Generalized Causal Inference. Houghton Mifflin Company, 2002. p105.
109. Rosenbaum PR, Rubin D. Reducing bias in observational studies using subclassification on the propensity score. Journal of the American Statistical Association. 1984; 79: 516-524.
110. Rosenbaum PR. Observational Studies. New York: Springer-Verlag. 1995.
111. Smith BD, Smith GL, Haffty BG. Postmastectomy radiation and mortality in women with T1-2 node-positive breast cancer. J Clin Oncol. 2005 Mar 1;23(7):1409-19.
112. Rice TW, Adelstein DJ, Chidel MA, et al. Benefit of postoperative adjuvant chemoradiotherapy in locoregionally advanced esophageal carcinoma. J Thorac Cardiovasc Surg. 2003 Nov;126(5):1590-6.
113. Krzyzanowska MK, Weeks JC, Earle CC. Treatment of locally advanced pancreatic cancer in the real world: population-based practices and effectiveness. J Clin Oncol. 2003 Sep 15;21(18):3409-14.
114. Indurkhya A, Mitra N, Schrag D. Using propensity scores to estimate the cost-effectiveness of medical therapies. Stat Med. 2005 Sep 12; [Epub ahead of print]
115. Shah, BR, Laupacis A, Hux JE, et al. Propensity score methods give similar results to traditional regression modeling in observational studies: a systematic review. J Clin Epid. 2005 58: 550-559.
116. Morriss, J. Personal communication.
117. Ontario Mortality Database Website.
http://www.healthinformation.on.ca/DataInfo/mortality.htm
118. Evans WK, Crook J, Read D, et al. Capturing tumour stage in a cancer information database. Cancer Prev Control. 1998 Dec;2(6):304-9.
119. Yau JC, Chan A, Eapen T, et al. Accuracy of the oncology patients information

system in a regional cancer centre. Oncol Rep. 2002 Jan-Feb;9(1):167-9.
120. Yau JC. Personal communication.
121. Martin BC, Ganguly R. SAS programming for Propensity Score Matching. http://www.rx.uga.edu/main/home/cas/faculty/propensity.pdf. Last accessed 7 June 2006.
122. Collet, D. Modelling Survival Data in Medical Research, Second edition. CRC Press, 2003. p130.
123. Harrell Jr, Frank E. Regression Modeling Strategies with applications to linear models, logistic regression, and survival analysis. Springer-Verlang, 2001. p24.
124. Shadish WR, Cook TD, Campbell DT. Experimental and Quasi-experimental Designs for Generalized Causal Inference. Houghton Mifflin Company, 2002. p163.

7.0 Supplementary Tables and Graphs

Table 3.1s. Relationship between categorical variables and adjuvant chemotherapy, all patients				
		Chemo		
Variable	Level	Yes	No	LRχ^2
T Stage	1A	6	128	<0.0001
	1B	48	394	
	1C	292	798	
	2	448	553	
	3	122	62	
	4A	13	14	
	4B	23	24	
	4C	3	10	
	4D	43	4	
Side	Left	477	1002	0.1747
	Right	521	985	
Grade	Low	111	543	<0.0001
	Intermediate or unknown	510	1255	
	High	377	189	
ER	Positive	577	1407	<0.0001
	Negative	369	244	
	Unknown	52	336	
PR	Positive	528	1249	<0.0001
	Negative	416	394	
	Unknown	54	344	
Nodal Status	Negative	340	1543	<0.0001
	Positive	658	444	
N Stage	0	323	1462	<0.0001
	1a	21	23	
	1bi	293	188	
	1bii	121	44	
	1biii	106	62	
	1biv	45	29	
	2	26	15	
	3	4	1	
	Unknown	29	173	
Year	1985	41	195	<0.0001
	1988	82	293	
	1992	140	383	
	1995	191	370	
	1998	235	404	
	2001	309	342	

Table 3.2s. Relationship between age at diagnosis and Adjuvant chemotherapy, all patients				
Model	**Variable**	**OR**	**Deviance**	**df**
1	Age	0.91	3032.094	1
2	Age rank1	1.00	3050.731	4
	Age rank2	0.550		
	Age rank3	0.176		
	Age rank4	0.088		
	Age rank5	0.029		

Table 3.3s. Relationship between tumour size and Adjuvant chemotherapy, all patients				
Model	**Variable**	**OR**	**Deviance**	**df**
1	Size	1.037	3503.033	1
2	Size rank1	1.000	3429.209	4
	Size rank2	2.817		
	Size rank3	5.315		
	Size rank4	6.384		
	Size rank5	12.760		
3	Ln(Size)	3.333	3409.907	1

Table 3. 4s. Relationship between categorical variables and adjuvant chemotherapy, node-positive patients				
		Chemo		
Variable	Level	Yes	No	LRχ^2
T Stage	1A	5	2	0.0237
	1B	33	25	
	1C	179	109	
	2	277	168	
	3	93	32	
	4A	8	6	
	4B	20	12	
	4C	3	4	
	4D	28	4	
Side	Left	344	188	0.6878
	Right	302	174	
Grade	Low	88	50	<0.0001
	Intermediate or Unknown	362	265	
	High	196	47	
ER	Pos	432	295	<0.0001
	Neg	179	33	
	Unkn	35	34	
PR	Pos	395	258	<0.0001
	Neg	214	68	
	Unkn	37	36	
N Stage	0	0	0	0.0097
	1a	21	23	
	1bi	293	188	
	1bii	121	44	
	1biii	106	62	
	1biv	75	29	
	2	26	15	
	3	4	1	
	Unk	0	0	
Stage	1	0	0	0.0492
	2A	212	135	
	2B	271	163	
	3A	101	38	
	3B	62	26	
Year	1985	37	55	<0.0001
	1988	74	73	
	1992	91	76	
	1995	109	75	
	1998	151	41	
	2001	184	42	

Table 3.5s. Relationship between age at diagnosis and chemo for node-positive patients				
Model	**Variable**	**OR**	**Deviance**	**df**
1	Age	0.868	844.260	1
2	Age rank1	1.00	866.081	4
	Age rank2	0.231		
	Age rank3	0.046		
	Age rank4	0.014		
	Age rank5	0.004		

Table 3.6s. Relationship between number of positive nodes and chemo for node-positive patients				
Model	**Variable**	**OR**	**Deviance**	**df**
1	Nodes	1.076	1301.763	1
2	nodes rank1	1.000	1302.677	4
	nodes rank2	1.119		
	nodes rank3	1.642		
	nodes rank4	1.730		

Table 3.7s. Relationship between tumour size and chemo for node-positive patients				
Model	**Variable**	**OR**	**Deviance**	**df**
1	Size	1.008	1307.982	1
2	Size rank1	1.000	1302.635	4
	Size rank2	1.120		
	Size rank3	0.827		
	Size rank4	1.050		
	Size rank5	1.728		

Table 3.8s. Relationship between categorical variables and adjuvant chemotherapy, node-negative patients				
		Chemo		
Variable	Level	Yes	No	LRχ^2
T Stage	1A	1	126	<0.0001
	1B	15	369	
	1C	113	689	
	2	171	385	
	3	29	30	
	4A	5	8	
	4B	3	12	
	4C	0	6	
	4D	15	0	
Side	Left	177	797	0.6736
	Right	175	828	
Grade	Low	23	493	<0.0001
	Intermediate or unknown	148	990	
	High	181	142	
ER	Positive	145	1112	<0.0001
	Negative	190	211	
	Unknown	17	302	
PR	Positive	133	991	<0.0001
	Negative	202	326	
	Unknown	17	308	
Year	1985	4	140	<0.0001
	1988	8	220	
	1992	49	307	
	1995	82	295	
	1998	84	363	
	2001	125	300	
Stage	1	129	1186	<0.0001
	2A	171	383	
	2B	28	29	
	3A	1	1	
	3B	23	26	

Table 3.9s. Relationship between age at diagnosis and chemo administration in node-negative patients				
Model	**Variable**	**OR**	**Deviance**	**df**
1	Diag_age	0.918	1552.740	1
2	Age rank1	1.000	1561.566	4
	Age rank2	0.593		
	Age rank3	0.198		
	Age rank4	0.084		
	Age rank5	0.044		

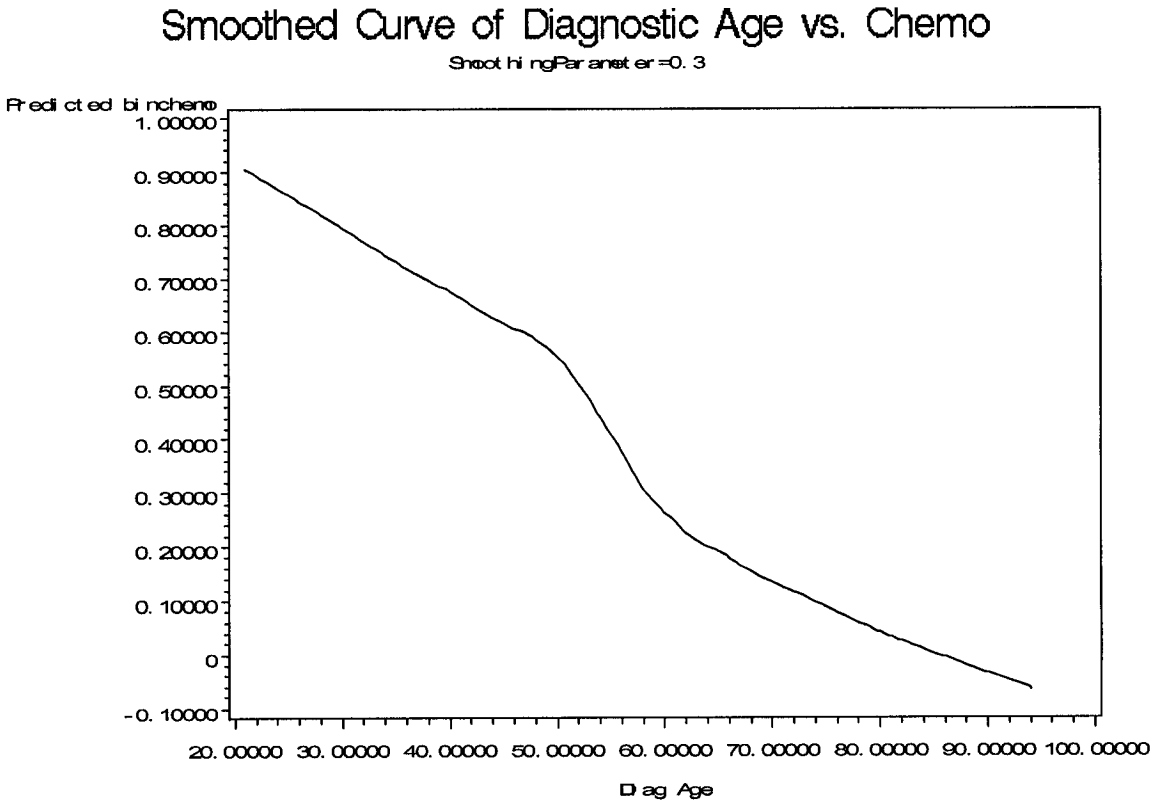
Table 3.10 s. Relationship between tumour size and chemo administration in node-negative patients				
Model	**Variable**	**OR**	**Deviance**	**df**
1	Size	1.045	1689.813	1
2	Size rank1	1.000	1633.782	4
	Size rank2	4.093		
	Size rank3	8.800		
	Size rank4	16.69		
	Size rank5	27.97		
3	ln(size)	4.302	1624.081	1
4	Size	1.091	1654.22	2
	$Size^2$	1.000		
5	Size	1.197	1626.641	3
	$Size^2$	0.997		
	$Size^3$	1.000		

Table 3.11s. Different formats of age for Cox regression			
Model Number	**Variable**	**HR**	**-2LogL**
1	Age	1.031	12032.041
2	Age rank1	1.000	11944.047
	Age Rank2	0.847	
	age rank3	1.018	
	Age rank4	1.369	
	Age rank5	2.756	
3	Age	1.029	11961.778
	Age^2	1.001	
	Age^3	1.000	
4	Age	1.028	11961.817
	Age^2	1.001	

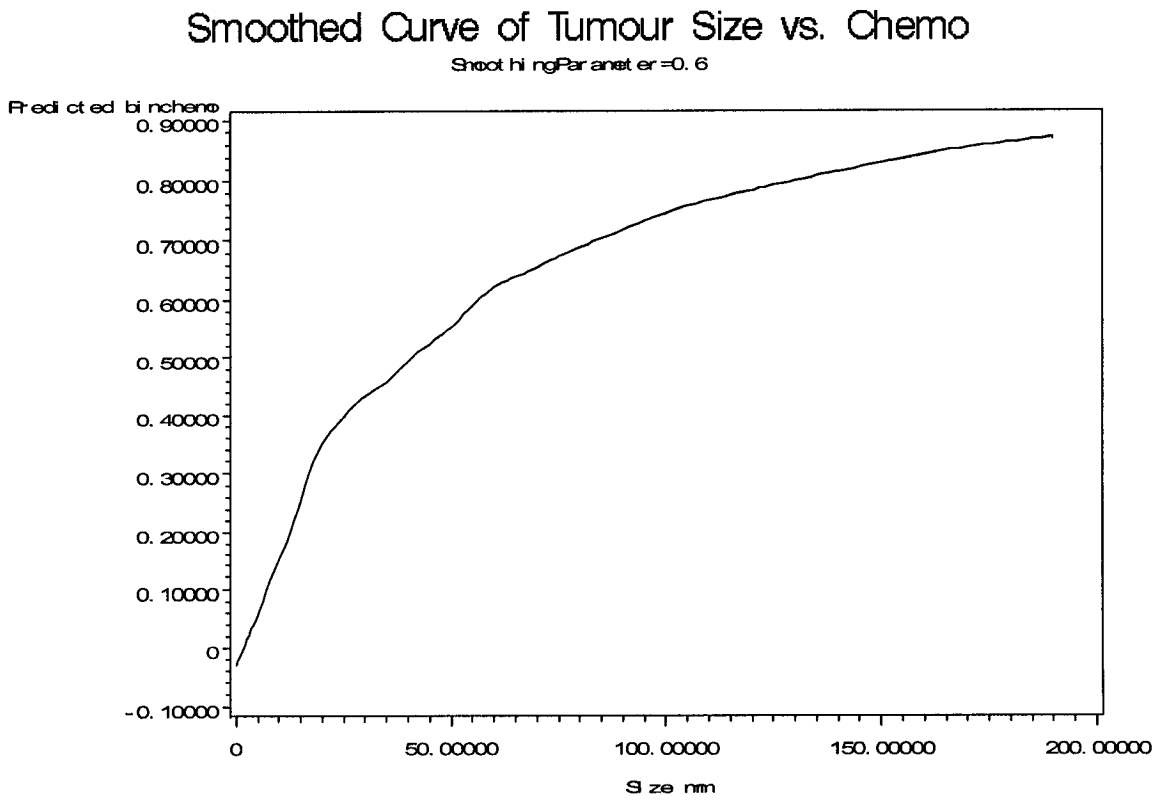
Table 3.12s. Different formats of tumour size for Cox regression			
Model Number	**Variable**	**HR**	**-2LogL**
1	Size_mm	1.017	12034.674
2	Size rank1	1.000	12035.247
	Size Rank2	1.551	
	Size rank3	1.513	
	Size rank4	2.239	
	Size rank5	3.149	

Table 3.13s. Different formats of number of nodes for Cox regression			
Model Number	**Variable**	**HR**	**-2LogL**
1	# of nodes	1.080	12092.372
2	Nodes rank1	1.000	12103.961
	Nodes Rank2	0.558	
	Nodes rank3	***	
	Nodes rank4	0.640	
	Nodes rank 5	****	
3	Sqrt # of nodes	1.344	12084.070

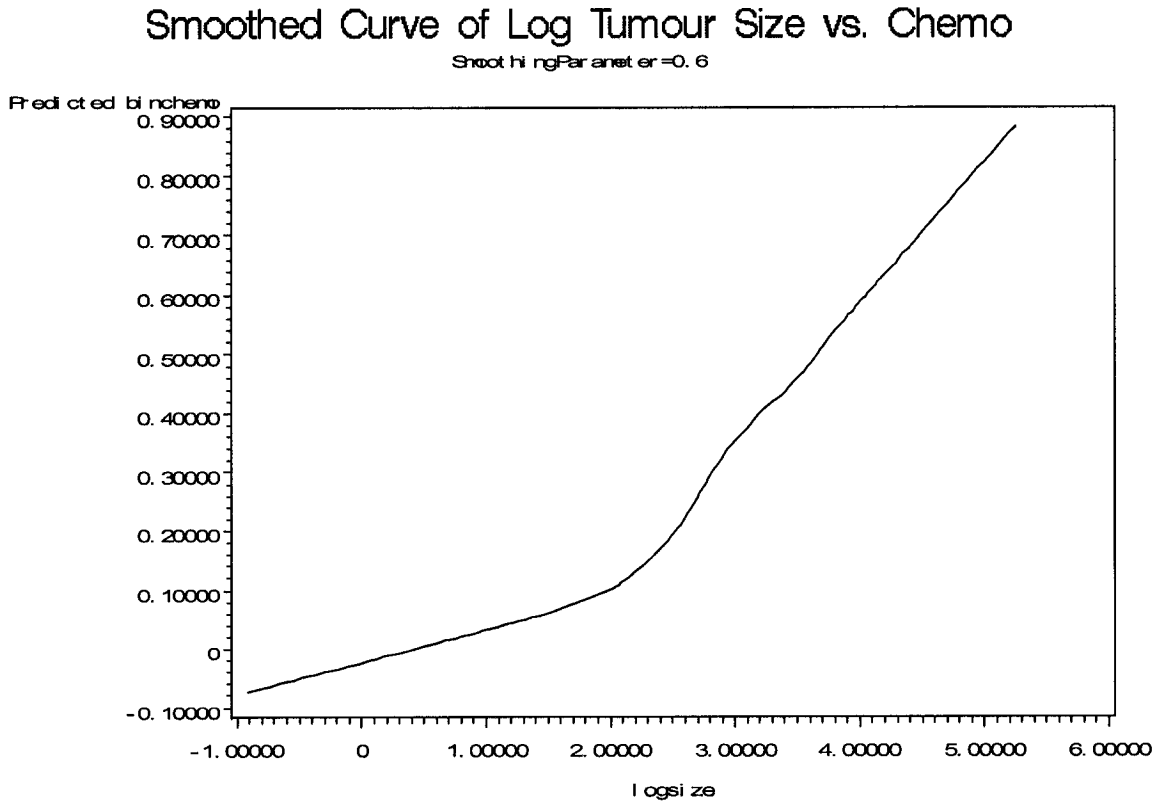
Graph 3.1s.



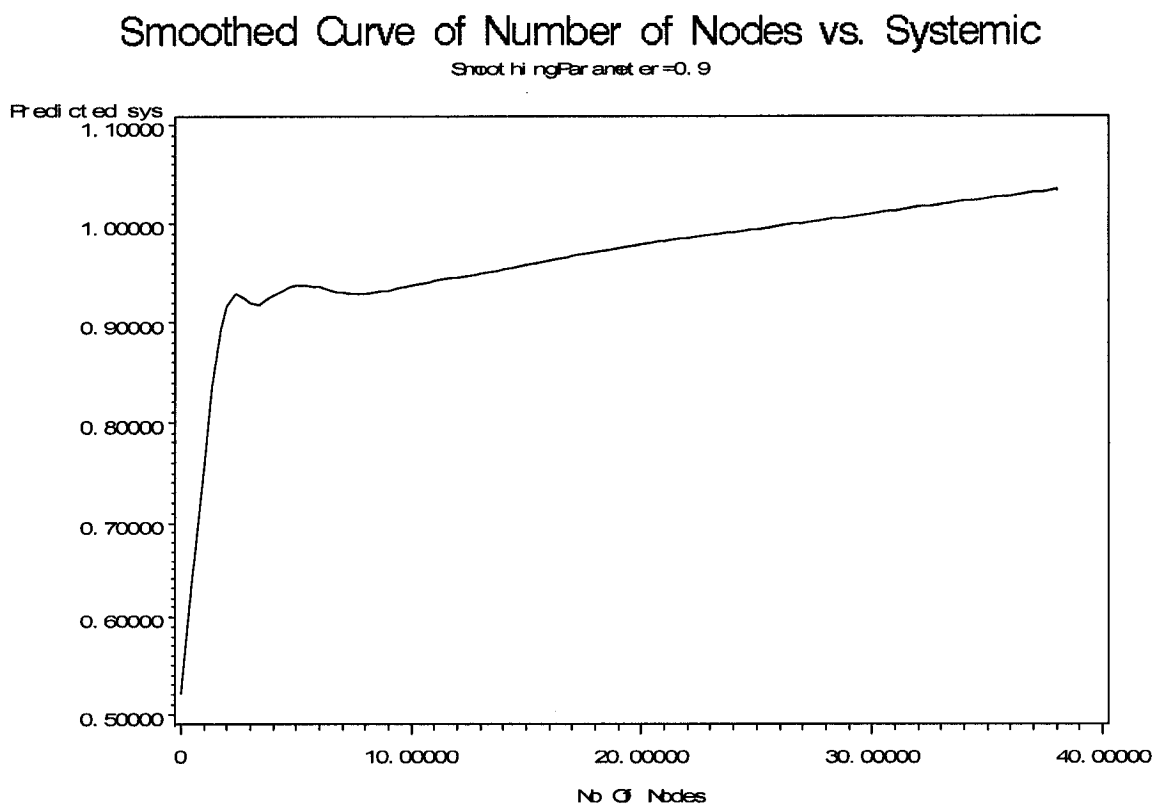
Graph 3.2s



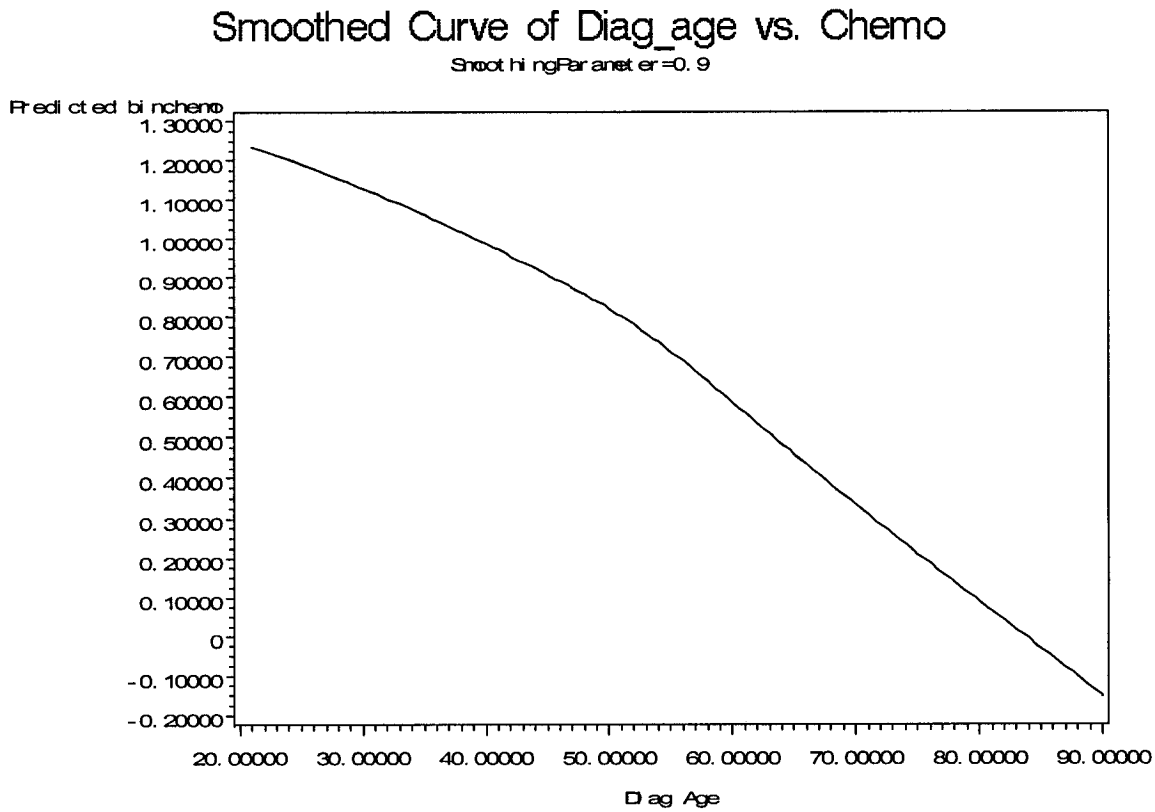
Graph 3.3s.



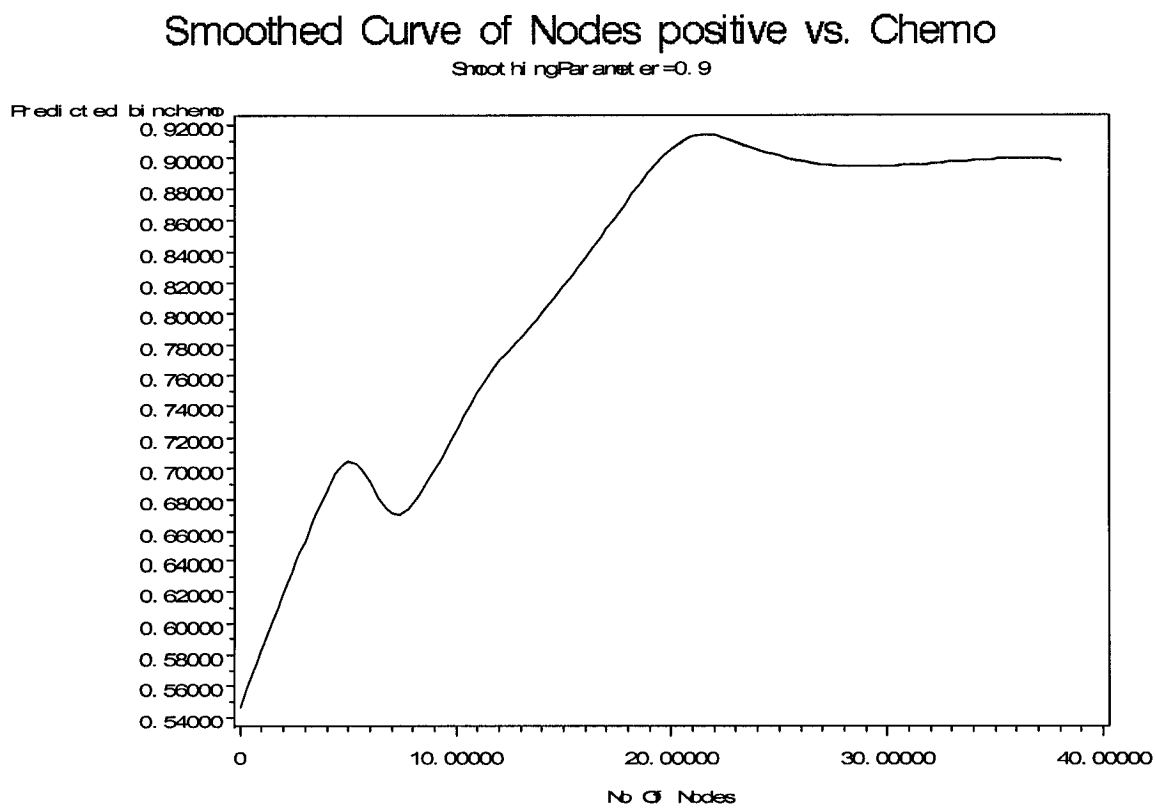
Graph 3.4s.



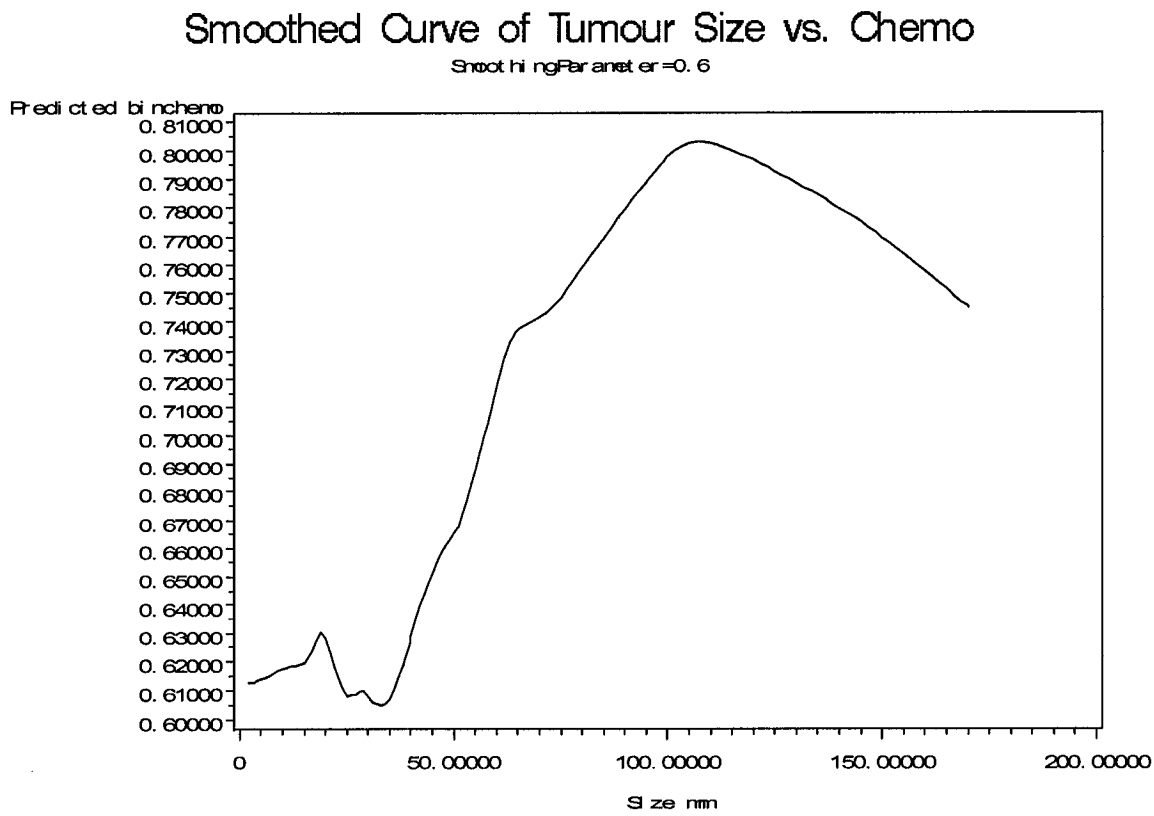
Graph 3.5s



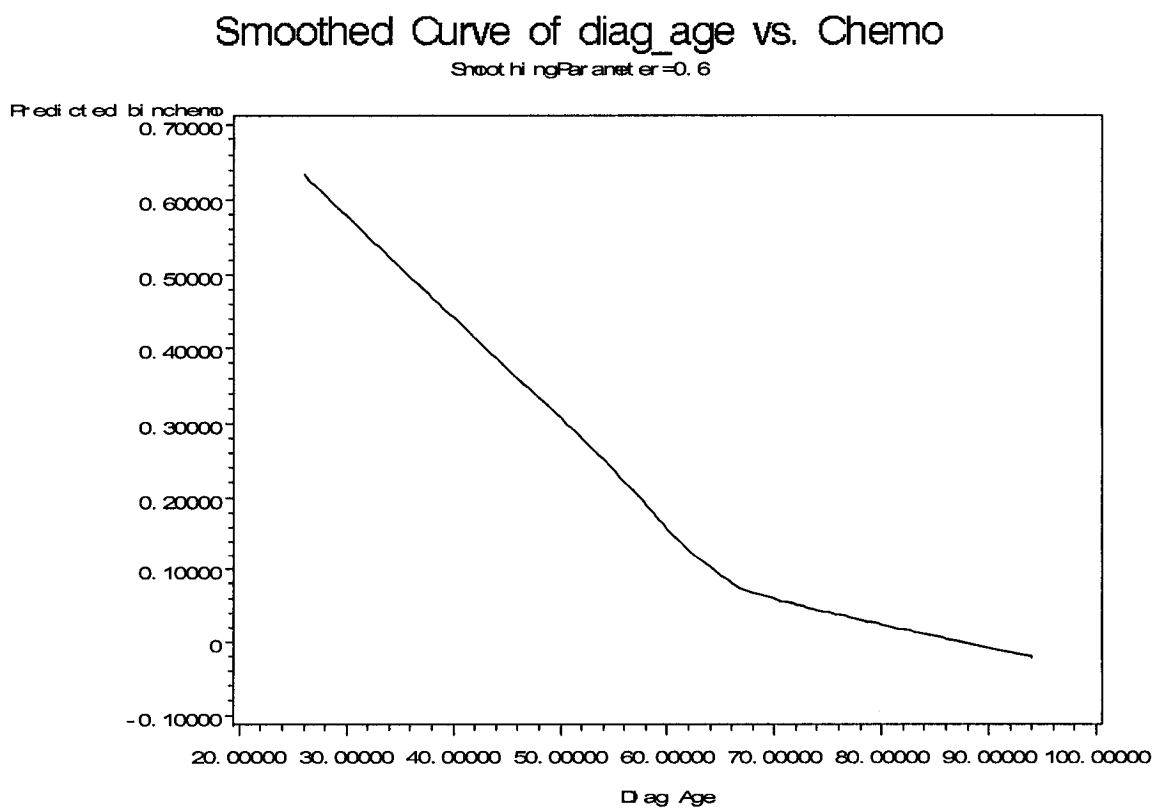
Graph 3.6s



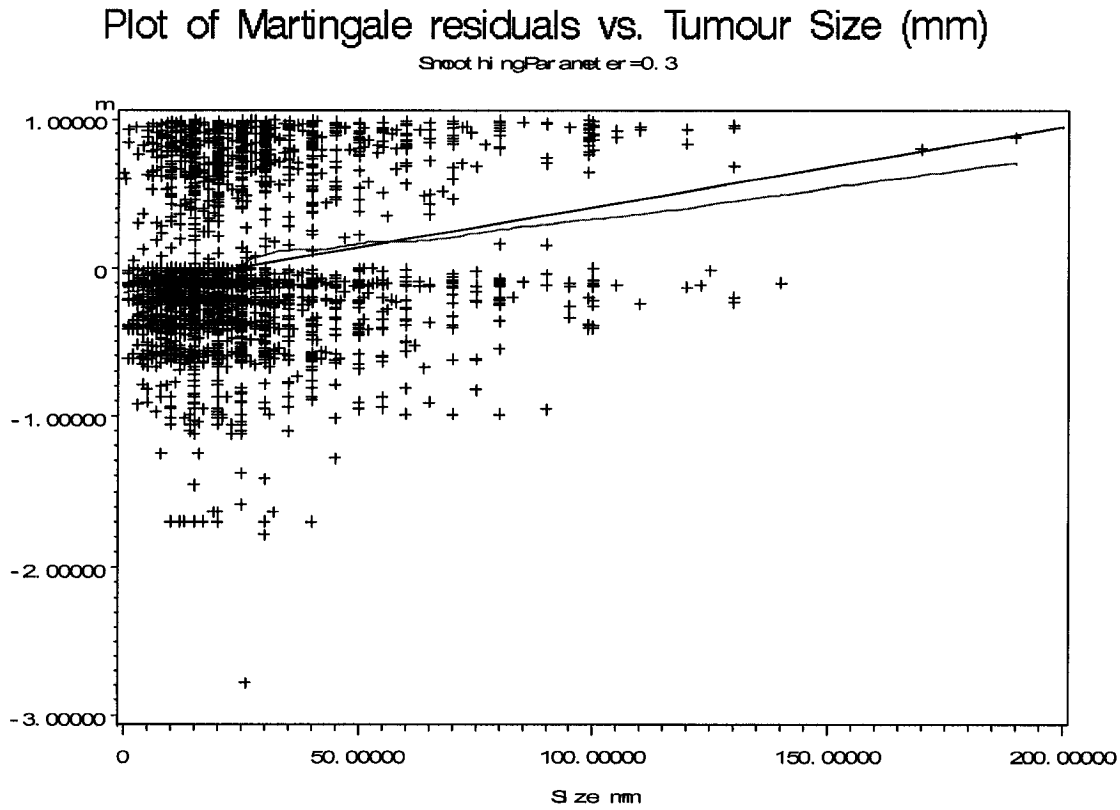
Graph 3.7s



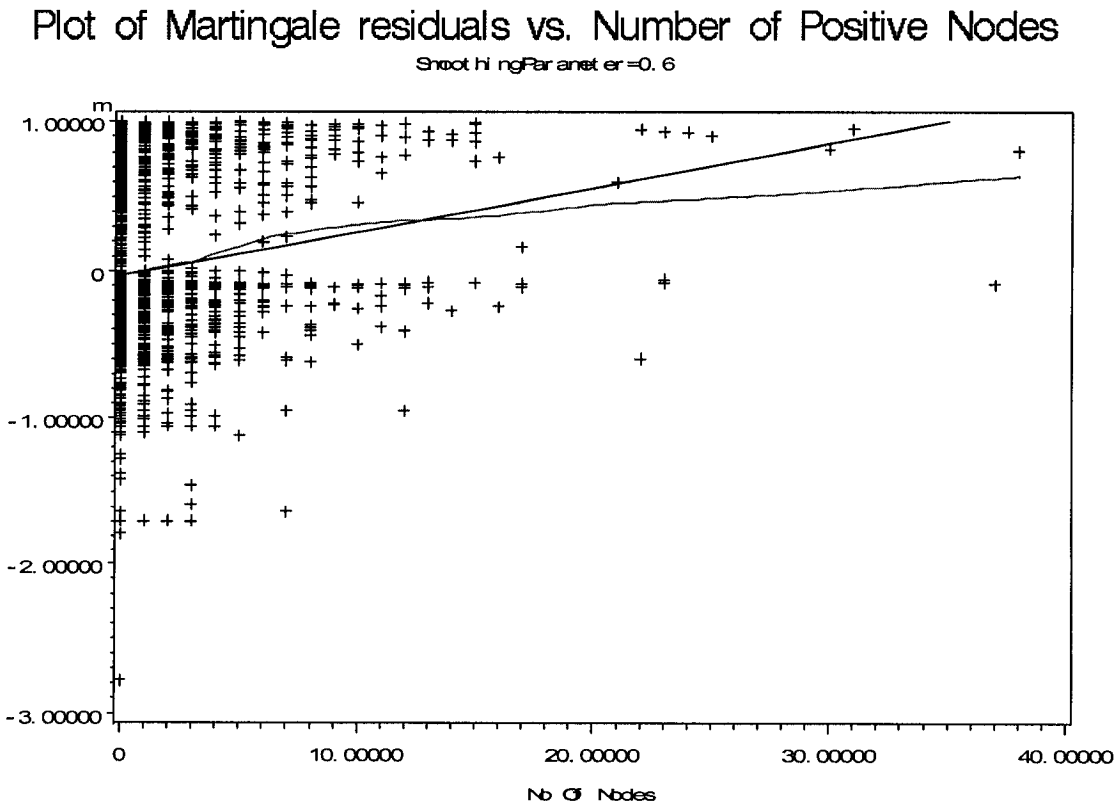
Graph 3.8s



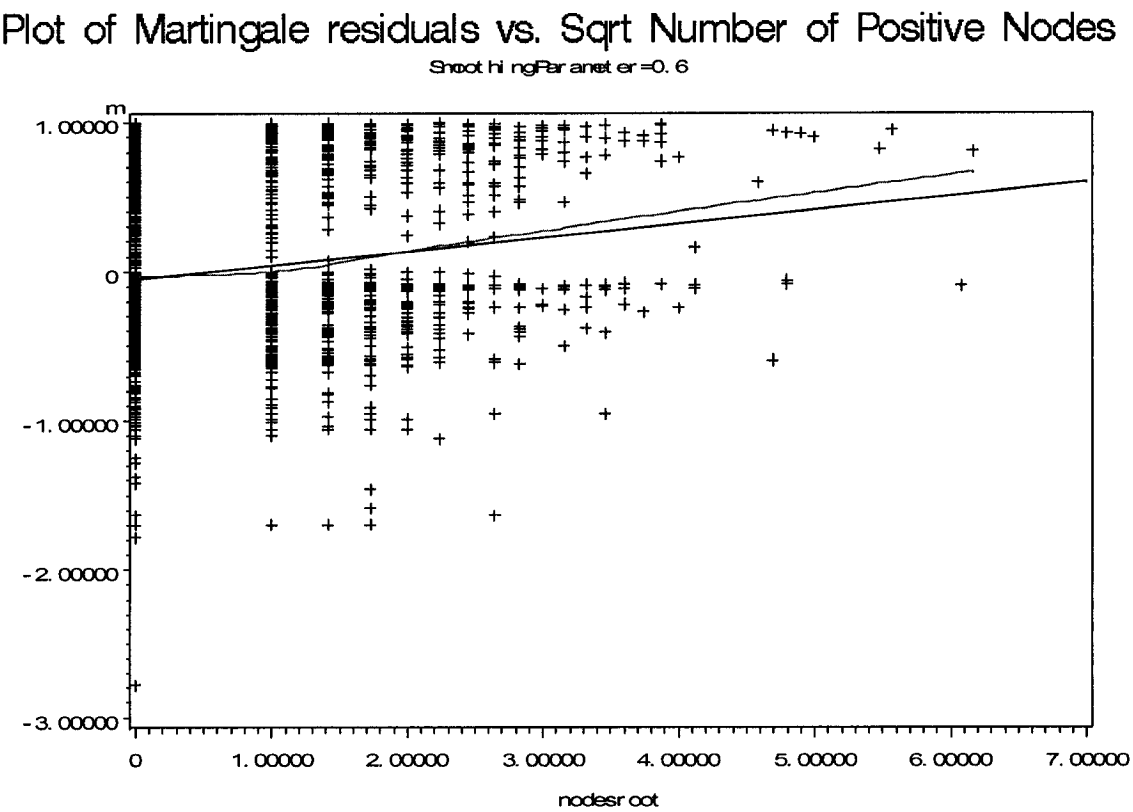
Graph 3.9s



Graph 3.10s



Graph 3.11s


Plot of Martingale residuals vs. Sqrt Number of Positive Nodes
Smoothing Parameter =0.6
m
1.00000
0
-1.00000
-2.00000
-3.00000
0
1.00000
2.00000
3.00000
4.00000
5.00000
6.00000
7.00000
nodesroot