UC Irvine
UC Irvine Previously Published Works

Title

Abstracting height and weight from medical records, and breast cancer pathologic factors

Permalink

https://escholarship.org/uc/item/0m8740kz

Journal

Cancer Causes & Control, 19(10)

ISSN

0957-5243

Authors

McEligot, Archana J
Im, Theresa
Dillman, Robert O
et al.

Publication Date

2008-12-01

DOI

10.1007/s10552-008-9192-z

Copyright Information

Peer reviewed

ORIGINAL PAPER

Abstracting height and weight from medical records, and breast cancer pathologic factors

Archana J. McEligot · Theresa Im · Robert O. Dillman ·
John G. West · Rana Salem · Reina Haque ·
Hoda Anton-Culver

Abstract Cancer registries routinely collect data on clinicopathologic factors, but rarely abstract anthropometric variables. We conducted a chart review study, examining the feasibility of abstracting weight, height, alcohol use, and smoking from medical records in women ($n = 1,974$) diagnosed with invasive breast cancer, and investigated the association between the abstracted variables with clinicopathologic features. Qualitative data were reviewed and categorized. Frequencies of the abstracted data, and demographic and clinicopathologic variables were calculated. Logistic regression models measured the relationship between the outcome variables, tumor size, stage of disease, and estrogen/progesterone (ER/PR) status with the abstracted variables. Data on current alcohol-use/no-use, current-smoker/non-smoker, and height/weight data were obtained on 96%, 97%, and 88–89% of the participants, respectively. The multivariate analysis showed that overweight (≥ 25 kg/m^2) women had significantly larger (≥ 2 cm) tumor size compared with normal weight for both women <50 years (OR = 1.79; 95% CI = 1.14–2.81; $p \leq 0.05$) and for women ≥ 50 years at diagnosis (OR = 1.58; 95% CI = 1.19–2.09; $p \leq 0.05$). These results suggest that abstracting current height and weight via medical records is feasible, and at minimum, current alcohol use and smoking status can be ascertained. In addition, being overweight was significantly associated with cancer clinicopathologic/prognostic factors, which has implications for monitoring etiologic factors that could be associated with cancer trends, incidence, and survival. Therefore routine collection of height and weight via cancer registries should be explored further.

Keywords Breast cancer · Obesity · Medical records · Clinical pathology

A. J. McEligot (✉)
Department of Health Science, California State University, Fullerton, 800 N. State College Blvd, KHS 254, Fullerton, CA 92834, USA
e-mail: amceligot@fullerton.edu

T. Im · H. Anton-Culver
Department of Epidemiology, University of California, Irvine, Irvine, CA, USA

R. O. Dillman
Hoag Cancer Center, Newport Beach, CA, USA

J. G. West
The Breast Care Center, Orange, CA, USA

R. Salem
Department of Rehabilitation Medicine, University of Washington Center on Outcomes Research in Rehabilitation, University of Washington, Box 354237, Seattle, WA 98195, USA
e-mail: RSalem@u.washington.edu

R. Haque
Research and Evaluation, Kaiser Permanente Medical Care Program Southern California, Pasadena, CA, USA

Introduction

Cancer registries regularly abstract data items pertaining to breast cancer clinicopathologic factors, including stage of disease, age at diagnosis, tumor size, estrogen receptor status, and several other factors. However, lifestyle and anthropometric factors including, height, weight, smoking, and alcohol use, which have been shown to be important etiologic and prognostic factors for breast cancer, are not routinely collected by cancer registries. The etiology of breast cancer and clinicopathologic factors, such as age at

diagnosis, tumor size, and estrogen/progesterone receptor (ER/PR) status, have been associated with height, weight, smoking, and alcohol use in some studies, but not in others [1–3].

Several studies have investigated the association between height, weight, and/or obesity with breast cancer risk [4–8]. Obesity, particularly for post-menopausal women, has been shown to increase risk (RR: 1.59, 95% CI: 1.09–2.32), while obesity in pre-menopausal women reduces risk (RR: 0.61, 95% CI: 0.42–0.89) [5, 7, 9]. Some studies also suggest that tumor size is related to obesity, specifically obese patients present with larger tumor size and/or diameter [1–3]. Maehle and colleagues found a significant association between larger tumor size (>2.0 cm) and body weight and body mass index (BMI) (2001). The relationship between obesity and tumor size was more pronounced in ER−/PR− tumors. However, other investigations have not found an association between body weight and clinicopathologic factors and therefore further studies need to be conducted [5, 10].

Other lifestyle factors, including smoking and alcohol use, have been associated with breast cancer clinicopathologic factors; however results have been equivocal, particularly for smoking. Some studies have suggested a weak association between smoking and poor prognosis after breast cancer diagnosis [11–13]. Manjer reported an increased risk of being diagnosed with ER− tumors [RR: 2.21 (95% CI: 1.23–3.96)] for current smokers and an increased risk of 2.67 (95% CI: 1.41–5.06) for ex-smokers compared with never smokers (2001). A recent study from the California Teachers cohort ($n = 116, 544$) indicated that current smokers have a 32% increased risk for breast cancer compared with never smokers, which is in contrast to previous studies that have shown no association between smoking and breast cancer either because passive smoking was not accounted for and/or the relationship only occurred in certain subgroups (younger women and/or women with family history) [14–18]. Alcohol use has also been associated with ER+/− status, with a recent study showing that alcohol users were significantly more likely (RR: 1.35, 95% CI 1.02–1.80) to be diagnosed with ER+ tumors, regardless of PR status, compared with non-drinkers in post-menopausal women [19]. But other studies have failed to show a relationship between alcohol use and estrogen receptor status [20–23].

Cancer prognosis has been divergent for non-Hispanic (NH) White women compared with other ethnicities [24–26]. Overall prognosis for breast cancer during the 1980s remained fairly steady, but after 1989 a dramatic decline in overall cancer mortality rates for non-Hispanic White women have been observed [26]. Furthermore, from 1980 to 1989 and more recently breast cancer mortality rates have declined 2.5% each year for White women [25, 27]. However, mortality for African American women

from 1973 to 1991 steadily increased, and has only recently slowed [24]. It is possible that access to care, genetic differences, and/or lifestyle differences may contribute to the observed differences in survival among various ethnic groups, but only one other study has investigated whether lifestyle differences, particularly height, weight, smoking, and alcohol use, could contribute to observed clinicopathologic differences (which may influence prognosis) between race/ethnic groups [28].

The present study assessed the feasibility and importance of routinely abstracting height and weight data by cancer registries. The purpose for the present study was to examine the feasibility of abstracting variables (height, weight, smoking, and alcohol use), not routinely collected by cancer registries, via medical chart review in women diagnosed with breast cancer. In addition, we investigated whether these factors are associated with clinicopathologic features of breast cancer, including tumor size, stage, tumor grade, and ER/PR status. A secondary aim was to further investigate the association between clinicopathologic characteristics and the abstracted items (height, weight, smoking, and alcohol use) by race/ethnicity.

Materials and methods

Study population and abstraction

We conducted chart reviews in 2003, and abstracted height, weight, smoking, and alcohol use data at time of diagnosis from existing medical records in a population-based cancer registry, the Cancer Surveillance Program of Orange County (CSPOC) [29–33]. Breast cancer cases diagnosed during a 1.5-year period between 1 July 1996 and 31 December 1997 were included in the present study. During this 1.5-year period, 2,275 cases were diagnosed with invasive breast cancer in Orange County. Of the 2,275 breast cancer cases, we were able to retrospectively abstract medical records on 1,974 patients. Therefore, abstraction was not completed on 301 patients due to missing charts, closed hospital and/or record facilities, or lack of on-site hospital resources (the hospital did not have either sufficient staff or space to accommodate our medical record abstractor).

Abstraction of height, weight, smoking, and alcohol use was conducted by an experienced medical records abstractor. The abstractor had previous experience in abstracting data from existing medical records charts, including IRB certification, which includes training in confidentiality of data and HIPPA certification. Data abstraction of height, weight, smoking, and alcohol use and date of data ascertainment from the chart were collected on a standardized form. Height in inches and weight in pounds were recorded. Smoking information, including current

smoking status, previous smoking history, cigarettes per day and years, and current and previous alcohol use, was also abstracted. Also, on a sub-sample of the patients where data was not available in the medical charts, we were able to obtain smoking and alcohol use status (that were collected for previous studies) from the registry database. In addition to abstracting quantitative data for smoking and alcohol use, we also recorded on a sub-sample of 25 patients (selected sequentially) qualitative data on smoking and alcohol use in order to more clearly characterize (quantity and frequency) these types of data in medical charts. We categorized the qualitative data for smoking as current smokers and non-smokers, and for alcohol use as current drinkers and non-drinkers. The number of charts abstracted per day was also recorded. The study protocol was approved by the Internal Review Board (IRB) of the University of California, Irvine (HSR#2003-3283), and by the California State University, Fullerton (HSR#07-0103).

Other study measures

Race/ethnicity and socioeconomic status (SES) data were ascertained via the cancer registry. SES score was based on various factors collected by the registry with a larger score indicating higher SES status [34]. Clinicopathologic features including stage, tumor size, tumor grade, ER/PR status, previous cancer history, and age at cancer diagnosis were also obtained through the cancer registry database. Stage of disease at diagnosis was the summary stage defined by the Surveillance, Epidemiology and End Results (SEER) program of the National Cancer Institute as follows: localized disease was defined as invasive carcinoma confined to the breast; regional stage was defined as invasive carcinoma spread beyond the breast, by direct extension and/or to regional lymph nodes; and distant disease was defined as direct extension beyond adjacent organs specified as regional, metastasis to distant lymph nodes, or development of discontinuous secondary or metastatic tumors. In terms of TNM classification, localized disease includes tumors T1–T3, N0, M0. Regional disease includes tumors T4, N0, M0 or any T, N1–N3, M0, and distant disease corresponds to any T, N, M1.

Statistical methods

We conducted statistics on aggregated patient data that were available for abstraction ($n = 1,974$). Height and weight data were used to calculate BMI (weight in kg/height in m^2). Frequencies were calculated on the following data items: age at diagnosis (categorized by 5 years), race/ethnicity [categorized as NH White, Asian/Pacific Islander (PI), Hispanic, NH Black, American Indian/Alaska Native, Other, and Unknown], BMI [categorized <18.5 (underweight), 18.5–24.9 (normal weight), 25.0–29.9 (overweight), 30.0+ (obese), Unknown], current smoking status (current smoker, non-smoker, unknown), and current alcohol use (current alcohol drinker, non-alcohol use, unknown). Additional frequencies on clinicopathologic features included: stage [localized, regional (direct extension, lymph nodes or ext and lymph), distant metastasis and unknown], tumor grade (grade I–IV, unknown), tumor size (<2 cm, 2+ cm, unknown), ER status (+/−, inconclusive, unknown), PR status (+/−, inconclusive, unknown), ER/PR status (ER+/PR+, ER+/PR−, ER−/PR+, ER−/PR−), prior cancer history (one cancer diagnosis only, two or more cancer diagnosis, unknown). Unknown data for the clinicopathologic features and race/ethnicity were due to missing information on these variables in the registry database.

We conducted logistic regression analysis to assess the relationship between the abstracted variables (BMI, smoking, and alcohol use) with prognostic variables, including tumor size, stage at diagnosis, and ER status. We excluded women diagnosed with more than one cancer ($n = 450$) and women who were underweight (<18.5 kg/m^2; $n = 46$) as this may indicate other existing co-morbid conditions. Three separate models for the outcome variables, tumor size (≥2 cm vs. <2 cm), stage at diagnosis (regional/distant vs. localized), and ER status (ER− vs. ER+), were conducted. Independent variables included in these models were: BMI (18.5–24.9 vs. 25.0+ kg/m^2), smoking (non-smoker vs. smoker) and alcohol use (use vs. no-use), age at diagnosis (continuous), race/ethnicity (NH White vs. other), and SES (continuous). All three models were stratified by age at diagnosis (<50 years and ≥50 years) in order to adjust for potential menopausal influences on these clinicopathologic factors. In addition, we conducted a chi-square analysis to assess the relationship between the abstracted variables and prognostic variables with race/ethnicity (NH White vs. other).

Results

Demographic, height and weight, and alcohol and smoking data are shown in Table 1. Twenty-six percent were diagnosed with breast cancer at age <50 years, while a majority of the women were diagnosed at age ≥50 years (74%). A majority of the women were NH White (85.9%) followed by Asian/PI (6.2%) and Hispanic (6.2%). 2.3% of the women were underweight, 42.6% were normal weight, 27.3% were overweight, and 15.9% were obese. A majority of the women were non-smokers (84.7%), while nearly half of the women were current alcohol users (46.4%).

Clinicopathologic features showed that 64.5% were diagnosed with localized stage, 29.8% had regional stage, and 3.7% had distant metastasis (Table 2). Tumor grade

Table 1 Distribution of demographic, BMI, alcohol use, and smoking data in women with invasive breast cancer (n = 1,974)

	n	%
Age at diagnosis		
24–29	13	0.7
30–39	125	6.3
40–49	376	19.0
50–59	418	21.2
60–69	413	20.9
70–79	400	20.3
80+	220	11.2
Unknown	9	0.5
Race/ethnicity		
Non-Hispanic (NH) White	1,695	85.9
Asian/PI	123	6.2
Hispanic	122	6.2
NH Black	17	0.9
American Indian/Alaskan Native	1	0.1
Other	3	0.2
Unknown	13	0.7
BMI (kg/m^2)		
<18.5 (underweight)	46	2.3
18.5–24.9 (normal)	841	42.6
25.0–29.9 (overweight)	539	27.3
30.0+ (obese)	313	15.9
Unknown	235	11.9
Current smoking status		
Non-smoker	1,671	84.7
Smoker	243	12.3
Missing/unknown	60	3.0
Current alcohol use		
No	972	46.4
Yes	915	49.2
Missing/unknown	87	4.4

Table 2 Distribution of clinicopathologic variables in women diagnosed with invasive breast cancer (n = 1,974)

	n	%
Stage		
Localized	1,274	64.5
Regional, direct extension	32	1.6
Regional, lymph nodes	499	25.3
Regional, ext and lymph	58	2.9
Distant metastasis	74	3.7
Unknown	37	1.9
Tumor grade		
Grade I, well diff.	413	20.9
Grade II, mod diff.	680	34.4
Grade III, poorly diff.	544	27.6
Grade IV, undiff.	41	2.1
Unknown	296	15.0
Tumor size		
<2 cm	1,060	53.7
2+ cm	796	40.3
Unknown	118	6.0
Estrogen (ER) status		
Positive	1,115	56.5
Negative	343	17.4
Inconclusive	4	0.2
Unknown	512	25.9
Progesterone (PR) status		
Positive	883	44.7
Negative	461	23.4
Inconclusive	17	0.9
Unknown	613	31.1
ER/PR status		
ER+/PR+	841	42.6
ER+/PR−	170	8.6
ER−/PR+	42	2.1
ER−/PR−	289	14.6
Unknown	632	32.0
Cancer history		
One cancer diagnosis only	1,515	77.1
Two or more cancer diagnoses	450	22.9
Unknown	9	0.5

was distributed as follows across grade I (20.9%), grade II (34.4%), and grade III (27.6), while 2.1% had grade IV disease. Approximately 54% were diagnosed with a tumor size of <2 cm, and 40.3% were diagnosed with a ≥2 cm tumor size. A majority had ER+ (56.5%) and PR+ (44.7%) tumors, and most had no history of previous cancers (76.7%).

Abstraction of height, weight, smoking, and alcohol use data from medical charts was conducted on six charts per hour. Height and weight abstraction was feasible on nearly 89% of the patients, and smoking and alcohol use was abstracted from medical records on 80.5% and 77.9%, respectively (Table 3). Height and weight data were found in multiple areas of the charts, usually in the anesthesiology sheet if the patient had surgery. Also, height and weight data were recorded in admission summary sheets,

and MD and chemotherapy notes (data not shown). With the inclusion of already existing data collected by the registry, smoking and alcohol use reached 97% and 96% completion (Table 3). The qualitative data on smoking and alcohol use shows that there is wide variation in recording smoking and alcohol use history with no standard method of recording years of smoking, number of cigarettes, and/or how often and how much a patient drinks alcohol in medical charts. Chart review indicated that qualitatively data for smoking history varied and included responses

Table 3 Feasibility of data collection, and completeness of quantitative and qualitative data ($n = 1,974$)

Abstracted variables	Frequency		Range of qualitative data in charts
	n	%	Sub-sample ($n = 25$)
Height (Feet and inches)			
Medical records	1,742	88.2	
Weight (Pounds)			
Medical records	1,751	88.7	
Current smoking status			Categorized as smoker / Categorized as non-smoker
Medical records	1,582	80.5	Smokes socially — No
			Somewhat less than a pack per day — No cigarettes
			Two packs/month × 20 years — Unknown or N/A
			30 pack years — None
			Smokes 1½ ppd — Never smoked
			Smokes 15 cigarettes per day for nine years — Does not use
			½ ppd for 40 years — Denies
			½ ppd for the last 20 years and continue to smoke — Denies any tobacco use
			1 ppd for over 20 years — Non-smoker
			1–2 ppd for a long time
			1.5 ppd × multiple years
			2 ppd/NOS
			2 ppd × 27 years
Registry Database	329	16.7	
Current alcohol use			Categorized as users / Categorized as non-users
Medical records	1,524	77.9	Used to drink one drink/day but quit in 1996 (Dx 7/96) — No
			Drinks alcohol — No does not drink
			Drinks socially — Unknown or N/A (no chart or from registry)
			Drinks rarely — None
			Drinks alcohol minimally — ETOH
			Rarely ever used alcohol — Does not use
			Modestly — Denies
			Drinks very little — Denies any alcohol use ever
			Special occasions — Non-ETOH
			Occasional alcohol user. Three to six per week — Non-alcoholic
			Occasional alcohol — Never
			Occasional glass of wine. A glass of wine as tonic daily — ETOH neg.
			On occasions drinks alcohol (wine w/dinner) — Not currently
			Occasional ETOH (two glasses of red wine per day) — No history of alcohol use
			Two glasses of wine, two to three times a week
			Less than one/week
			Occasional ETOH (one/month)
			Rarely uses alcohol. On social occasions. Once or twice a week
			Less than one drink a day
Registry Database	354	18.1	

such as smokes socially, somewhat less than a pack per day, 30 pack-years, 2 packs/month × 20, and smokes 1.5 ppd. Alcohol use was recorded in the charts as drinks alcohol, drinks socially, and drinks rarely. Therefore when collecting smoking and alcohol use from medical records,

smoking and alcohol use can only be categorized into current smoker/non-smoker and current use/no alcohol use.

Multivariate logistic regression (Table 4), adjusted for age at diagnosis, race/ethnicity, and SES, showed that being overweight was significantly (OR: 1.79, 95%

Table 4 Logistic regression models[a,b] of clinicopathologic outcomes associated with BMI, smoking and alcohol use

	Age at diagnosis <50 years		Age at diagnosis ≥50 years	
	OR	(95% CI)	OR	(95% CI)
Tumor size (≥2 cm vs. <2 cm)	$n = 360$		$n = 861$	
BMI 18.5–24.9	1.00	Ref	1.00	Ref
BMI 25.0+	1.79[c]	(1.14–2.81)	1.58[c]	(1.19–2.09)
Non-smoker	1.00	Ref	1.00	Ref
Smoker	0.56	(0.30–1.05)	1.26	(0.81–1.96)
No alcohol	1.00	Ref	1.00	Ref
Alcohol	1.12	(0.70–1.76)	0.90	(0.67–1.19)
Stage (regional/met vs. localized)	$n = 288$		$n = 676$	
BMI 18.5–24.9	1.00	Ref	1.00	Ref
BMI 25.0+	0.93	(0.57–1.53)	0.94	(0.67–1.32)
Non-smoker	1.00	Ref	1.00	Ref
Smoker	1.44	(0.75–2.79)	0.81	(0.49–1.33)
No alcohol	1.00	Ref	1.00	Ref
Alcohol	1.29	(0.79–2.12)	1.29	(0.91–1.83)
ER status (negative vs. positive)	$n = 289$		$n = 679$	
BMI 18.5–24.9	1.00	Ref	1.00	Ref
BMI 25.0+	0.83	(0.49–1.41)	0.96	(0.66–1.46)
Non-smoker	1.00	Ref	1.00	Ref
Smoker	0.88	(0.44–1.78)	0.61	(0.31–1.12)
No alcohol	1.00	Ref	1.00	Ref
Alcohol	0.96	(0.57–1.63)	0.58[c]	(0.39–0.88)

[a] Adjusted for age at diagnosis, race/ethnicity, and socioeconomic status (SES)

[b] Excludes those with BMI < 18.5 kg/m^2 and those with two or more cancers

[c] $p < 0.05$

CI: 1.14–2.81, $p \leq 0.05$) associated with larger (≥2 cm) tumor size in women <50 years at age of diagnosis. Similarly in women ≥50 years at diagnosis, the odds of a larger tumor size was 1.58 (95% CI: 1.19–2.09, $p \leq 0.05$) in women who were overweight compared with women who were normal weight (18.5–24.9 kg/m^2). Also, alcohol drinkers were significantly ($p < 0.05$) less likely (OR: 0.58, 95% CI: 0.39–0.88) to be ER− compared to non-drinkers, for women ≥50 years at diagnosis. No relationship was observed between PR+/− status and alcohol use (data not shown).

Table 5 shows the clinicopathologic and lifestyle variable proportions by race/ethnicity. When assessing the clinicopathologic features stratified by race/ethnicity, the data suggests that a significantly higher proportion of the other ethnic group (all women other than NH White) were diagnosed at a younger age ($p < 0.0001$), were diagnosed with regional/metastatic disease ($p = 0.07$), have larger tumor size ($p = 0.005$), have higher tumor grade ($p = 0.005$), and have a higher proportion of ER− and PR− tumor status ($p < 0.0001$), compared with NH White women. When examining the abstracted lifestyle/anthropometric variables, a lower proportion of the other ethnic

group smoked ($p = 0.09$), drank less ($p < 0.001$), were shorter ($p < 0.001$) and were more overweight, but not significantly ($p = 0.12$), compared with the NH White group.

Discussion

In the present study, abstraction of height and weight data was feasible on nearly 90% of the sample and we obtained current smoking and alcohol use data at time of diagnosis on approximately 97% of the sample from medical charts and registry data. In terms of BMI, approximately half were normal weight and the other half were either overweight or obese. A majority of the women were non-smokers and half (46%) drank alcohol. Younger women (<50 years at diagnosis) who were overweight were 80% more likely to be diagnosed with larger tumors compared with women who were normal weight. Similarly, women diagnosed at age ≥50 years and who were overweight were 58% more likely to be diagnosed with larger tumors. In the same age group, women who drank were 42% less likely to be diagnosed with ER− tumors (vs. ER+ tumors)

Table 5 Distribution[a] of abstracted and clinicopathologic variables stratified by NH White and other race/ethnicity

	Race/ethnicity				p-Value
	NH White		Other		
	n	%	n	%	
Abstracted variables					
BMI (kg/m^2)					
18.5–24.9	591	52.6	83	46.4	
25.0–29.9	347	30.9	63	35.2	
≥30.0	186	16.5	33	18.4	0.12
Total	1,124		179		
Height (in)					
<64	458	40.7	128	71.5	
64–66	445	39.6	41	22.9	
67+	221	19.7	10	5.6	<0.0001
Total	1,124		179		
Smoking					
Current	149	13	16	9	
Non-smoker	959	87	163	91	0.09
Total	1,108		179		
Alcohol use					
Current	589	54	52	29	
No use	508	46	127	71	<0.0001
Total	1,097		179		
Clinicopathologic variables					
Age at diagnosis (years)					
<40	63	5.6	27	15.1	
40–49	230	20.5	61	34.1	
≥50	831	73.9	91	50.8	<0.0001
Total	1,124		179		
Stage					
Localized	739	66.6	104	58.1	
Regional	337	30.4	67	37.4	
Distant metastasis	33	3.0	8	4.5	0.07
Total	1,109		179		
Tumor size					
<2 cm	619	57.4	79	45.9	
2+ cm	459	42.6	93	54.1	0.005
Total	1,078		172		
Tumor grade					
Grade I/II	646	66.9	87	55.4	
Grade III/IV	320	33.1	70	44.6	0.005
Total	966		157		
Estrogen (ER) status					
Positive	675	78.7	84	62.7	
Negative	183	21.3	50	37.3	<0.0001
Total	858		134		

Table 5 continued

	Race/ethnicity				p-Value
	NH White		Other		
	n	%	n	%	
Progesterone (PR) status					
Positive	539	67.8	63	54.3	
Negative	256	32.2	53	45.7	0.004
Total	795		116		
ER/PR status					
ER+/PR+	512	64.5	59	50.9	
ER+/PR−	106	13.4	10	8.6	
ER−/PR+	27	3.4	4	3.4	
ER−/PR−	149	18.8	43	37.1	0.0001
Total	794		116		

[a] Excludes cases who have been diagnosed with two or more cancers and/or BMI < 18.5

compared with non-drinkers. In addition, we found that NH White women have overall better clinicopathologic markers/status at diagnosis compared with other ethnic groups, but the other ethnic groups were more likely to be overweight and less likely to drink or smoke.

Cancer registries routinely collect clinicopathologic features, including stage of disease, age at diagnosis, tumor size, and tumor grade; however few studies have assessed the feasibility of abstracting modifiable risk factors, including weight, smoking, and alcohol use, from medical records. One earlier study that assessed obesity and breast cancer recurrence and survival in African-American women reported that height, weight, ER/PR status, tumor size, and other prognostic factors were not routinely found in medical records [35]. In contrast, for the present study, abstraction of height and weight data reached nearly 90% and with previous data from the registry, smoking and alcohol use data was abstracted on nearly 100% of the charts. It is highly possible that because there was a six-year difference from when we began abstracting to the actual date of diagnosis (we reviewed charts in 2003 for women diagnosed between 1996 and 1997) that we were not able to abstract height and weight data on nearly all patients due to missing charts at storage facilities and hospital closures. Our results also indicate that at minimum we were only able to collect whether patients were current smokers and/or alcohol users, because more specific details on number of cigarettes, pack-years, or type/frequency of alcohol used were not consistently recorded in medical charts.

Several studies have shown that being overweight and/or obese increases the risk for breast cancer in post-

menopausal women by 20% [36–38]. Our results showed that being overweight increased the risk of being diagnosed with a larger tumor size in both younger (<50 years) and older women (≥50 years). Similar to our results, a population-based study of 1,177 women younger than 45 years (pre-menopausal) showed that women in the highest BMI quartile had a significantly larger tumor size (2 to <5 cm: OR, 2.3; 95% CI, 1.5–3.1; or ≥5 cm: OR, 2.7; 95% CI, 1.5–4.8) compared with women in the lowest BMI quartile [39]. Other studies have shown a relationship between higher BMI and larger tumor size in post-menopausal women alone [3, 38]. The Iowa Women's Health Study revealed that post-menopausal women in the highest BMI tertile were more likely to be diagnosed with a larger tumor (≥2 cm) compared with women in the lowest tertile [40]. Ove Maehle and colleagues followed up women enrolled in the Norwegian Cancer Registry and showed that patients who were overweight (26.2 kg/m^2) had a significantly larger tumor diameter (≥2 cm) compared with women with a lower BMI (25.2 kg/m^2) [3]. The relationship between tumor size and BMI was significantly present only in women >50 years and in ER negative and PR negative tumors. It is possible that the relationship between tumor size and weight could be mediated by hormonal fluctuations in women who are overweight, and studies have suggested this [3, 33]; however, similar to our results, another study showed no relationship between obesity and/or BMI with ER/PR status [11].

In our study, alcohol use, a modifiable risk factor, was associated with a reduced risk of ER negative tumors when compared with ER positive tumors in post-menopausal women, indicating a relationship between alcohol use and ER positive status. A population-based study in 1,188 breast cancer patients showed that post-menopausal women in the highest alcohol intake strata had a 35% increased risk (RR: 1.35, 95% CI: 1.56–3.56) of being diagnosed with ER positive tumors [19]. A recent cohort study with a follow-up of 10 years in 38, 454 women suggested that women drinkers were at increased risk (RR: 1.11, 95% CI: 1.03–1.20) for breast cancer for ER positive/PR positive tumors only, while no relationship was found with alcohol use in women diagnosed with ER negative tumors [41]. Both hormonal and non-hormonal dependent mechanisms have been suggested in relation to alcohol use and breast cancer risk [42–44]. For the present study, it is possible that we found a relationship between alcohol use and ER positive status, because alcohol increases estrogen production and also increases the expression of ER positive tumors [43, 44].

Our results showed that NH White women compared with other ethnic groups (includes Hispanic, Asian, NH Black, and other) had better clinicopathologic indicators, including lower stage, tumor size, tumor grade at diagnosis, and were more likely to be ER positive, PR positive, and ER/PR positive. In contrast to the NH White women, the other ethnic group had better lifestyle behaviors, including current smoking and drinking less. However, the other ethnic group had higher BMI compared with NH White women (not statistically significant). Several other studies have suggested differences in breast cancer incidence and outcomes among various ethnic groups [28, 45, 46]. However, few studies have compared modifiable risk factors and clinicopathologic factors among ethnic groups. The Women's Health Initiative followed 156,570 women and compared breast cancer characteristics in 3,938 women among six ethnic groups and reported that compared to White women, African American and Hispanic women were significantly less likely ($p < 0.001$) to be diagnosed with ER positive and PR positive tumors, and had significantly higher frequency of poorly differentiated tumors [28]. However, no significant differences in tumor size and stage were reported among the five ethnic groups. Similar to our results, the same study reported that compared to Whites, the other ethnic groups, except for Asian/Pacific Islander, had significantly higher BMI, and women of every minority group were significantly ($p < 0.001$) less likely to drink alcohol compared to White women. It is widely recognized that early screening through mammography may reduce differences in stage and other clinicopathologic factors among the various ethnic groups [47, 48]. But, these studies did not adjust for BMI and/or other modifiable factors when assessing differences in clinicopathologic factors among the ethnic groups.

Limitations of the present study should be recognized. First, abstraction of alcohol use and smoking data via medical charts revealed that, at most, only never and/or current use data for these modifiable factors was available. Number of cigarettes smoked, pack years, and alcohol drinks per day and type of alcohol were not quantifiable via medical charts. Therefore, associations between alcohol use and smoking status with clinicopathologic factors were limited to use and/or no use of current alcohol and smoking. Second, even though we abstracted variables at time of diagnosis, because there was a six-year difference from when the participants were diagnosed (in 1996) to when we conducted the chart review (in 2003), some hospitals had closed and/or charts were missing and therefore height and weight data were obtained on only 89% of patients. However, based on abstraction of alcohol and smoking data (which included collection via the registry for previous studies), if height and weight data were abstracted along with the standard clinicopathologic variables, then most likely nearly 100% of height and weight data could be ascertained. Last, we did not consistently record the location of the abstracted variables in the medical charts which could affect feasibility of collecting these data. However, obtaining the precise location of all the abstracted variables

in the charts was beyond the scope of the present study and should be conducted in future investigations.

There are several strengths of the present study. The data obtained for the present study was via a population-based cancer registry, and includes an ethnically diverse population. Additionally, weight data was measured and then recorded in the anesthesiology records/medical charts, as well as the clinicopathologic variables, which provide objective data. Compared to self-reported data for weight which might be subject to socially desirable responses, objective measures reduce the chances of spurious findings. Importantly, this is one of the first studies to assess the feasibility of abstracting height and weight data via chart review.

The continuing increasing trend in obesity in the United States and the relationship between being overweight and/ or obese with increased cancer incidence should emphasize the importance of collecting measurements of height and weight via cancer registries, making this a public health priority. The results of the present study indicate that abstracting height and weight data from medical charts is feasible, and can be conducted in a relatively short period of time. However, smoking and alcohol use is not consistently recorded and may be difficult to report systematically. Also, these data suggest that BMI (and the other lifestyle variables) are associated with clinicopathologic factors at breast cancer diagnosis, and that these modifiable risk factors do vary by race/ethnicity. Because registries routinely collect important clinicopathologic factors and follow-up data associated with cancer diagnosis, adding routine abstraction of height and weight data may increase our understanding of changes in cancer incidence trends, disparities among ethnic groups, and etiologic factors associated with cancer. Therefore, cancer registries should, at least preliminarily, begin to explore collection of height and weight data, and further investigate implementation of routine abstraction of these two data items that have significant public health impact and association with cancer risk.

References

1. Baumgartner KB, Hunt WC, Baumgartner RN et al (2004) Association of body composition and weight history with breast cancer prognostic markers: divergent pattern for Hispanic and non-Hispanic White women. Am J Epidemiol 160:1087–1097. doi:10.1093/aje/kwh313

2. Loi S, Milne RL, Friedlander ML et al (2005) Obesity and outcomes in premenopausal and postmenopausal breast cancer. Cancer Epidemiol Biomarkers Prev 14:1686–1691. doi:10.1158/1055-9965.EPI-05-0042

3. Maehle BO, Tretli S, Skjaerven R, Thorsen T (2001) Premorbid body weight and its relations to primary tumour diameter in breast cancer patients; its dependence on estrogen and progesteron receptor status. Breast Cancer Res Treat 68:159–169. doi:10.1023/A:1011977118921

4. Friedenreich CM (2001) Review of anthropometric factors and breast cancer risk. Eur J Cancer Prev 10:15–32. doi:10.1097/00008469-200102000-00003

5. Hunter DJ, Willet WC (1993) Diet, body size, and breast cancer. Epidemiol Rev 15:110–132

6. Negri E, La Vecchia C, Bruzzi P et al (1988) Risk factors for breast cancer: pooled results from three Italian case–control studies. Am J Epidemiol 128:1207–1215

7. Tehard B, Clavel-Chapelon F (2006) Several anthropometric measurements and breast cancer risk: results of the E3N cohort study. Int J Obes (Lond) 30:156–163. doi:10.1038/sj.ijo.0803133

8. van den Brandt PA, Spiegelman D, Yaun SS et al (2000) Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. Am J Epidemiol 152:514–527. doi:10.1093/aje/152.6.514

9. Huang Z, Hankinson SE, Colditz GA et al (1997) Dual effects of weight and weight gain on breast cancer risk. JAMA 278:1407–1411. doi:10.1001/jama.278.17.1407

10. Ursin G, Longnecker MP, Haile RW, Greenland S (1995) A meta-analysis of body mass index and risk of premenopausal breast cancer. Epidemiology 6:137–41. doi:10.1097/00001648-199503000-00009

11. Daniell HW (1988) Increased lymph node metastases at mastectomy for breast cancer associated with host obesity, cigarette smoking, age, and large tumor size. Cancer 62:429–435. doi:10.1002/1097-0142(19880715)62:2<429::AID-CNCR2820620230>3.0.CO;2-4

12. Manjer J, Malina J, Berglund G et al (2001) Smoking associated with hormone receptor negative breast cancer. Int J Cancer 91:580–584. doi:10.1002/1097-0215(200002)9999:9999<::AID-IJC1091>3.0.CO;2-V

13. Moritz DJ, Satariano WA (1993) Factors predicting stage of breast cancer at diagnosis in middle aged and elderly women: the role of living arrangements. J Clin Epidemiol 46:443–454. doi:10.1016/0895-4356(93)90021-R

14. Brownson RC, Blackwell CW, Pearson DK, Reynolds RD, Richens JW, Papermaster BW (1988) Risk of breast cancer in relation to cigarette smoking. Arch Intern Med 148:140–144. doi:10.1001/archinte.148.1.140

15. Chu SY, Stroup NE, Wingo PA, Lee NC, Peterson HB, Gwinn ML (1990) Cigarette smoking and the risk of breast cancer. Am J Epidemiol 131:244–253

16. Couch FJ, Cerhan JR, Vierkant RA et al (2001) Cigarette smoking increases risk for breast cancer in high-risk breast cancer families. Cancer Epidemiol Biomarkers Prev 10:327–332

17. Reynolds P, Hurley S, Goldberg DE et al (2004) Active smoking, household passive smoking, and breast cancer: evidence from the California Teachers Study. J Natl Cancer Inst 96:29–37

18. Terry PD, Rohan TE (2002) Cigarette smoking and the risk of breast cancer in women: a review of the literature. Cancer Epidemiol Biomarkers Prev 11:953–971

19. Suzuki R, Ye W, Rylander-Rudqvist T, Saji S, Colditz GA, Wolk A (2005) Alcohol and postmenopausal breast cancer risk defined by estrogen and progesterone receptor status: a prospective cohort study. J Natl Cancer Inst 97:1601–1608

20. Cooper JA, Rohan TE, Cant EL, Horsfall DJ, Tilley WD (1989) Risk factors for breast cancer by oestrogen receptor status: a population-based case–control study. Br J Cancer 59:119–125

21. Cotterchio M, Kreiger N, Theis B, Sloan M, Bahl S (2003) Hormonal factors and the risk of breast cancer according to estrogen- and progesterone-receptor subgroup. Cancer Epidemiol Biomarkers Prev 12:1053–1060

22. Huang WY, Newman B, Millikan RC, Schell MJ, Hulka BS, Moorman PG (2000) Hormone-related factors and risk of breast cancer in relation to estrogen receptor and progesterone receptor status. Am J Epidemiol 151:703–714

23. Yoo KY, Tajima K, Miura S et al (1997) Breast cancer risk factors according to combined estrogen and progesterone receptor status: a case–control analysis. Am J Epidemiol 146:307–314

24. Howe HL, Wingo PA, Thun MJ et al (2001) Annual report to the nation on the status of cancer (1973 through 1998), featuring cancers with recent increasing trends. J Natl Cancer Inst 93:824–842. doi:10.1093/jnci/93.11.824

25. MMWR (2002) Recent trends in mortality rates for four major cancers, by sex and race/ethnicity—United States, 1990–1998. From the Centers for Disease Control and Prevention 287:1391–1392

26. Ries LAG, Eisner MP, Kosary CL et al (2000) SEER cancer statistics review, 1973–1997. National Cancer Institute, Bethesda, MD

27. Smigal C, Jemal A, Ward E et al (2006) Trends in breast cancer by race and ethnicity: update 2006. CA Cancer J Clin 56:168–183

28. Chlebowski RT, Chen Z, Anderson GL et al (2005) Ethnicity and breast cancer: factors influencing differences in incidence and outcome. J Natl Cancer Inst 97:439–448

29. Anton-Culver H, Culver BD, Kurosaki T, Osann KE, Lee JB (1988) Incidence of lung cancer by histological type from a population-based registry. Cancer Res 48:6580–6583

30. Seiffert JE, Price WT, Gordon B (1990) The California tumor registry: a state-of-the-art model for a regionalized, automated, population-based registry. Top Health Rec Manage 11:59–73

31. Anton-Culver H, Cohen PF, Gildea ME, Ziogas A (2000) Characteristics of BRCA1 mutations in a population-based case series of breast and ovarian cancer. Eur J Cancer 36:1200–1208. doi:10.1016/S0959-8049(00)00110-6

32. Ziogas A, Gildea M, Cohen P et al (2000) Cancer risk estimates for family members of a population-based family registry for breast and ovarian cancer. Cancer Epidemiol Biomarkers Prev 9:103–111

33. Verreault R, Brisson J, Deschenes L, Naud F (1989) Body weight and prognostic indicators in breast cancer. Modifying effect of estrogen receptors. Am J Epidemiol 129:260–268

34. Yost K, Perkins C, Cohen R, Morris C, Wright W (2001) Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. Cancer Causes Control 12:703–711. doi:10.1023/A:1011240019516

35. Watzlaf VJ, Katoh A, D'Amico F (1996) Obstacles encountered in the use of the medical record and cancer registry abstract in breast cancer research. Top Health Inf Manage 17:25–33

36. Carmichael A (2006) Obesity as a risk factor for development and poor prognosis of breast cancer. BJOG 113:1160–1166. doi:10.1111/j.1471-0528.2006.01021.x

37. La Vecchia C, Negri E, Franceschi S, Talamini R, Bruzzi P, Palli D, Decarli A (1997) Body mass index and post-menopausal breast cancer: an age-specific analysis. Br J Cancer 75:441–444

38. Petrelli JM, Calle EE, Rodriguez C, Thun MJ (2002) Body mass index, height, and postmenopausal breast cancer mortality in a prospective cohort of US women. Cancer Causes Control 13:325–332. doi:10.1023/A:1015288615472

39. Daling JR, Malone KE, Doody DR, Johnson LG, Gralow JR, Porter PL (2001) Relation of body mass index to tumor markers and survival among young women with invasive ductal breast carcinoma. Cancer 92:720–729. doi:10.1002/1097-0142(20010815)92:4<720::AID-CNCR1375>3.0.CO;2-T

40. Zhang S, Folsom AR, Sellers TA, Kushi LH, Potter JD (1995) Better breast cancer survival for postmenopausal women who are less overweight and eat less fat. The Iowa Women's Health Study. Cancer 76:275–283. doi:10.1002/1097-0142(19950715)76:2<275::AID-CNCR2820760218>3.0.CO;2-6

41. Zhang SM, Lee IM, Manson JE, Cook NR, Willett WC, Buring JE (2007) Alcohol consumption and breast cancer risk in the Women's Health Study. Am J Epidemiol 165:667–676. doi:10.1093/aje/kwk054

42. Brooks PJ (1997) DNA damage, DNA repair, and alcohol toxicity—a review. Alcohol Clin Exp Res 21:1073–1082

43. Fan S, Meng Q, Gao B et al (2000) Alcohol stimulates estrogen receptor signaling in human breast cancer cell lines. Cancer Res 60:5635–5639

44. Singletary KW, Gapstur SM (2001) Alcohol and breast cancer: review of epidemiologic and experimental evidence and potential mechanisms. JAMA 286:2143–2151. doi:10.1001/jama.286.17.2143

45. Lantz PM, Mujahid M, Schwartz K et al (2006) The influence of race, ethnicity, and individual socioeconomic factors on breast cancer stage at diagnosis. Am J Public Health 96:2173–2178. doi:10.2105/AJPH.2005.072132

46. Sassi F, Luft HS, Guadagnoli E (2006) Reducing racial/ethnic disparities in female breast cancer: screening rates and stage at diagnosis. Am J Public Health 96:2165–2172. doi:10.2105/AJPH.2005.071761

47. McCarthy EP, Burns RB, Freund KM et al (2000) Mammography use, breast cancer stage at diagnosis, and survival among older women. J Am Geriatr Soc 48:1226–1233

48. Oluwole SF, Ali AO, Adu A et al (2003) Impact of a cancer screening program on breast cancer stage at diagnosis in a medically underserved urban community. J Am Coll Surg 196:180–188. doi:10.1016/S1072-7515(02)01765-9