

The full version of this article
may be viewed online at
jop.ascopubs.org

Stanford University School of Medicine, Stanford; Palo Alto Medical Foundation Research Institute, Palo Alto; San Diego State University, San Diego; Cancer Prevention Institute of California, Fremont; Genomic Health Inc, Redwood City, CA; and Geisel School of Medicine, Lebanon, NH

Corresponding author: Allison W. Kurian, MD, MSc, Stanford University School of Medicine, HRP Redwood Building, Rm T254A, 150 Governor's Lane, Stanford, CA 94305-5405; e-mail: akurian@stanford.edu.

Disclosures provided by the authors are available with this article at jop.ascopubs.org.

Use of Gene Expression Profiling and Chemotherapy in Early-Stage Breast Cancer: A Study of Linked Electronic Medical Records, Cancer Registry Data, and Genomic Data Across Two Health Care Systems

Anosheh Afghahi, MD, MPH, Maya Mathur, MS, Caroline A. Thompson, PhD, Aya Mitani, MPH, Joseph Rigdon, PhD, Manisha Desai, PhD, Peter P. Yu, MD, Monique A. de Bruin, MD, Tina Seto, MS, Cliff Olson, MS, Pragati Kenkare, MS, Scarlett L. Gomez, PhD, Amar K. Das, MD, PhD, Harold S. Luft, PhD, George W. Sledge, Jr, MD, Amy P. Sing, MD, and Allison W. Kurian, MD, MSc

QUESTION ASKED: Do patients with breast cancer who seek treatment in more than one health care system receive more aggressive chemotherapy, even if a quantitative measure of their prognosis (eg, the 21-gene recurrence score [RS]) does not differ from that of other patients with breast cancer?

SUMMARY ANSWER: We identified a subset of patients for whom RS is used more often: patients treated in more than one health care system. These dual use patients differed from the larger group primarily by their higher socioeconomic status and not by any clinically relevant characteristics such as cancer prognostic factors or median RS. Reassuringly, RS-receipt patients had no differential use of chemotherapy associated with sociodemographic characteristics that were not pertinent to their cancer care. We speculate that emerging genomic technologies used to tailor medical therapies might help to optimize cancer treatment across health care settings.

WHAT WE DID: The aim of the Oncoshare project, a multisource database for breast cancer outcomes research, is to integrate data from electronic medical records (EMRs) from Stanford University Hospital (an academic medical center; hereafter University) and Palo Alto Medical Foundation (a multisite community health care system; hereafter Community) in the same geographic region. To provide a gold standard for patient identification and treatment summaries, Oncoshare uses patient-level linkage of EMR data to the population-based California Cancer Registry, a component of the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program. We extracted clinical data from the EMR systems. California Cancer Registry records were requested for all patients with breast cancer diagnosed and/or treated at either health care system. Eligible patients consisted of all women diagnosed with stage I to II, hormone receptor–positive, human epidermal growth factor receptor 2–negative breast cancer between 2005 and 2011. The administrative database of Genomic Health Inc, the laboratory that provides RS, was searched for all patients who had RS ordered as a component of cancer care. Multivariable analysis was performed to identify predictors of RS and chemotherapy use.

WHAT WE FOUND: A total of 10,125 patients with breast cancer were diagnosed in the University or Community systems from 2005 to 2011: 2,418 (23.9%) met RS guidelines criteria, of whom 15.6% received RS. RS was less often used for patients with involved lymph nodes, higher tumor grade, and age < 40 or ≥ 65 years. Among RS recipients, chemotherapy receipt was associated with a higher RS score (intermediate v low: odds ratio, 3.66; 95% CI, 1.94 to 6.91). Two hundred ninety-three patients (10.6%) received care in both health care systems (dual use); although RS receipt was associated with dual use (v University: odds ratio, 1.73; 95% CI, 1.18 to 2.55), there was no difference in chemotherapy use after RS by health care setting.



DOI: [10.1200/JOP.2015.009803](https://doi.org/10.1200/JOP.2015.009803);
published online ahead of print at
jop.ascopubs.org on May 24, 2016.

BIAS, CONFOUNDING FACTOR(S), DRAWBACKS: The Oncoshare database draws from a single catchment area in Northern California, focusing on two health care systems that disproportionately represent non-Hispanic whites and Asians of high socioeconomic status; thus, results may not be broadly generalizable. As with all observational studies, causality inferred from associations can be misleading as a result of uncontrolled confounding or selection biases. Reassuringly, we did not find any evidence of a difference in risk profile between the patients who received the RS versus those who did not, yet there may be unmeasured prognostic differences.

REAL-LIFE IMPLICATIONS: Our prior work using this integrated approach has consistently identified a cluster of apparent overuse of diagnostic and treatment interventions by patients accessing care across health care systems; here, we tested the hypothesis that such use patterns would extend to chemotherapy after genomic profiling. However, our results suggest that instead, ongoing efforts toward precision medicine, such as the RS, may reduce unwarranted variation and disparities in cancer care. This may be a potential mechanism by which patients who more often seek additional health care resources have their health care use mitigated by reassuring results of objective genomic testing. **JOP**

Apply for Quality Oncology Practice Initiative (QOPI) Certification to Join More Than 260 Certified Practices

Certification signifies that the practice or institution has:

- A commitment to provide quality cancer care
- The resources, programs, and services sufficient to accomplish and sustain that commitment
- Participated in a voluntary comprehensive site assessment against clearly specified national standards of care
- Succeeded in meeting the standards and objectives of the Certification Program

Learn more at qopi.asco.org/certification.html.



Use of Gene Expression Profiling and Chemotherapy in Early-Stage Breast Cancer: A Study of Linked Electronic Medical Records, Cancer Registry Data, and Genomic Data Across Two Health Care Systems

Anosheh Afghahi, MD, MPH, Maya Mathur, MS, Caroline A. Thompson, PhD, Aya Mitani, MPH, Joseph Rigdon, PhD, Manisha Desai, PhD, Peter P. Yu, MD, Monique A. de Bruin, MD, Tina Seto, MS, Cliff Olson, MS, Pragati Kenkare, MS, Scarlett L. Gomez, PhD, Amar K. Das, MD, PhD, Harold S. Luft, PhD, George W. Sledge, Jr, MD, Amy P. Sing, MD, and Allison W. Kurian, MD, MSc

Stanford University School of Medicine, Stanford; Palo Alto Medical Foundation Research Institute, Palo Alto; San Diego State University, San Diego; Cancer Prevention Institute of California, Fremont; Genomic Health Inc, Redwood City, CA; and Geisel School of Medicine, Lebanon, NH

Abstract

Purpose

The 21-gene recurrence score (RS) identifies patients with breast cancer who derive little benefit from chemotherapy; it may reduce unwarranted variability in the use of chemotherapy. We tested whether the use of RS seems to guide chemotherapy receipt across different cancer care settings.

Methods

We developed a retrospective cohort of patients with breast cancer by using electronic medical record data from Stanford University (hereafter University) and Palo Alto Medical Foundation (hereafter Community) linked with demographic and staging data from the California Cancer Registry and RS results from the testing laboratory (Genomic Health Inc., Redwood City, CA). Multivariable analysis was performed to identify predictors of RS and chemotherapy use.

Results

In all, 10,125 patients with breast cancer were diagnosed in the University or Community systems from 2005 to 2011; 2,418 (23.9%) met RS guidelines criteria, of whom 15.6% received RS. RS was less often used for patients with involved lymph nodes, higher tumor grade, and age < 40 or ≥ 65 years. Among RS recipients, chemotherapy receipt was associated with a higher score (intermediate v low: odds ratio, 3.66; 95% CI, 1.94 to 6.91). A total of 293 patients (10.6%) received care in both health care systems (hereafter dual use); although receipt of RS was associated with dual use (v University: odds ratio, 1.73; 95% CI, 1.18 to 2.55), there was no difference in use of chemotherapy after RS by health care setting.

Conclusion

Although there was greater use of RS for patients who sought care in more than one health care setting, use of chemotherapy followed RS guidance in University and Community health care systems. These results suggest that precision medicine may help optimize cancer treatment across health care settings.

ASSOCIATED CONTENT



Appendix DOI: [10.1200/JOP.2015.009803](https://doi.org/10.1200/JOP.2015.009803)



DOI: [10.1200/JOP.2015.009803](https://doi.org/10.1200/JOP.2015.009803); published online ahead of print at jop.ascopubs.org on May 24, 2016.

INTRODUCTION

Most newly diagnosed breast cancers in the United States are hormone receptor–positive (HR–positive, defined as estrogen receptor–positive and/or progesterone receptor–positive), human epidermal growth factor receptor 2 (HER2)–negative, and axillary lymph node–negative.¹ Decisions about adjuvant chemotherapy are guided by the risk:benefit ratio of the increased probability of cure versus adverse effects of treatment. In addition to the pathologic features of tumors such as size, grade, and proliferation index, gene expression profiling with the 21-gene recurrence score (RS; *Oncotype DX Breast Cancer Assay*, Genomic Health Inc. [GHI], Redwood City, CA) among eligible patients (early-stage, HR-positive, HER2-negative disease) may improve risk stratification and target chemotherapy more effectively. Since RS use has been incorporated into national guidelines, its use has increased substantially, with 20% to 45% of eligible patients receiving the test and with an associated reduction in use of chemotherapy over time. RS has been considered cost-effective in its ability to target chemotherapy effectively.²⁻¹⁴

Studies of the use and associated outcomes of emerging cancer genomic tests have been limited by the quality of available data. State cancer registries offer demographics and staging data but lack detailed treatment information. Electronic medical records (EMRs) are often missing the results of send-out genomic tests, which are frequently returned to ordering clinicians in paper or PDF formats that may be scanned into EMRs but are rarely entered as structured data. Moreover, EMRs, which are institutionally dependent, often miss the complete spectrum of care received by patients who may seek treatment from multiple facilities. To address this gap, we developed Oncoshare, an integrated breast cancer database that links demographic and staging data from the population-based California Cancer Registry (CCR) with treatment information from the EMRs at Stanford University Hospital (hereafter University) and the Palo Alto Medical Foundation (hereafter Community), a nearby multisite community practice.¹⁵⁻¹⁷ We previously reported on the apparent greater use of health care by patients treated in both health care systems, including substantially more bilateral mastectomies and breast magnetic resonance imaging scans.¹⁵

On the basis of our prior findings, our hypothesis was that patients with breast cancer who were seeking treatment in more than one health care system would receive more aggressive adjuvant chemotherapy, even if a quantitative

measure of their prognosis (eg, the RS) did not differ from that of other patients with breast cancer. To test this hypothesis, we obtained RS results directly from the testing laboratory (GHI) and integrated them into the Oncoshare database by linkage at the patient level. A further aim was to characterize the demographics, use of health care systems, and cancer prognostic factors among those who received RS versus those who did not and to determine whether RS seemed to guide the use of chemotherapy across different health care settings.

METHODS

Data Resource

The Oncoshare project, a multisource database for research on breast cancer outcomes, was initiated in 2009; it aims to integrate data from the EMRs of an academic medical center (Stanford University Hospital [hereafter University]) and a neighboring multisite community health care system (Palo Alto Medical Foundation [hereafter Community]) in the same geographic region in the San Francisco Bay Area of Northern California. Although Community physicians occasionally provide inpatient care in University facilities, the institutions are legally and financially separate and do not have overlapping staff. In previous analyses, we identified a distinct group of patients who used services at both University and Community sites.^{15,17}

To provide a gold standard for identifying patients and treatment summaries, Oncoshare links patient-level EMR data to the population-based CCR. The CCR is a component of the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program and captures data on approximately 99% of all cancer cases diagnosed in California. The methods involved in developing Oncoshare, including validation of the data linkage strategy, have been published previously.^{15,16} Briefly, we extracted clinical data (treatment and diagnostic test data) from the University and Community EMR systems. Chemotherapy data elements in each EMR were mapped to a standardized drug lexicon (RxNorm)¹⁸ to ensure uniform coding. CCR records were requested for all patients with breast cancer diagnosed and/or treated at the University and/or Community. Data fields from the CCR included age, race/ethnicity, tumor stage, tumor grade, histology, HR and HER2 status, and treatment summaries from any treating California institution, not just the two systems participating in this study.

Census block groups were determined by geocoding patients' residential addresses at the time of diagnosis. Neighborhood socioeconomic status (SES) was assigned by using a previously developed index that incorporated data from the 2000 US Census (including education, income, occupation, and housing costs) on the basis of selection using principal components analysis.¹⁹ CCR and EMR records were linked by using patients' names, Social Security numbers, medical record numbers, and birth dates. To protect patient privacy, all protected health information was removed, and all calendar dates were shifted by a random factor of ± 30 days, with the shift factor and direction constant for each patient to maintain internal consistency.¹⁶ All research was approved by University, Community, and California State institutional review boards (the latter for use of CCR data).

Patient Identification and Variable Definition

Eligible patients consisted of all women diagnosed and/or treated by the University or Community health care systems with stage I to II HR-positive HER2-negative breast cancer between January 1, 2005 (when RS became available), through December 31, 2011 (the last year of complete follow-up available from CCR at the time of analysis). All patients were identified through the Oncoshare database, and thus all received breast cancer care in one or both of the participating health care systems. Diagnostic, staging, and HR data were obtained from the CCR, and treatment data were obtained from the CCR and EMRs. We did not restrict the sample to patients with node-negative breast cancer (for whom National Comprehensive Cancer Network [NCCN] evidence-based clinical practice guidelines currently recommend RS testing),²⁰ because recent data suggest a role for RS testing in node-positive breast cancer.²¹

The GHI administrative database was searched for all patients who had RS ordered as a component of cancer care by a provider affiliated with the University and/or Community health care system from the time of initial availability of the RS test in 2005 to the study cutoff date of December 2011. The RS yields a discrete numeric score from 0 to 100; prespecified cut points defined during test development were used to classify tumors as low risk (< 18), intermediate risk (18 to 30), or high risk (> 30), with risk scores predicting chemotherapy benefit (eg, low benefit with low score).⁹ A numeric RS value from 0 to 100 was provided by GHI for each patient identified as having received the test. Patients were classified as having received

adjuvant chemotherapy if CCR and/or EMR data from either health care system indicated that cytotoxic, antineoplastic medications had been initiated within 1 year of breast cancer diagnosis.

By using a previously published algorithm,¹⁷ we classified use of health care systems according to care periods and locations. For care periods, we included the diagnosis period, which was defined as 90 days before the diagnosis date until the date of first treatment; the treatment period, which started from the first intervention and concluded after a gap in recorded cancer care of at least 180 days; and any time during the cancer episode. For the care location, we included three mutually exclusive categories: University only, Community only, or both health care systems (dual use).¹⁷

Data Analysis

We defined two overlapping analytic cohorts: RS-eligible patients who met study eligibility criteria and were diagnosed within the University or Community health care systems, and RS-receipt patients who met study eligibility criteria and received RS under the care of a University or Community physician. We described our study population, stratified by cohort definition, by using frequencies and proportions for categorical variables and medians and interquartile ranges or ranges for continuous variables.

We used multivariable logistic regression techniques with a nested outcomes model in the appropriate cohorts: RS use among RS-eligible patients, chemotherapy use among RS-eligible patients, and chemotherapy use among RS-receipt patients. All models included the following covariates: age (< 40 , 40-49, 50-64, or ≥ 65 years), calendar year of diagnosis (2005 to 2007, 2008 to 2009, or 2010 to 2011), race (non-Hispanic white, non-Hispanic black, Hispanic, Asian/Pacific Islander, or other), neighborhood SES (in statewide quintiles), insurance status (private, Medicare, or other), cancer stage and grade, lymph node involvement, histology, and laterality.

Primary analyses incorporated multiple imputation (MI) methods to handle missing data.²²⁻²⁶ To perform a sensitivity analysis, we also fit models on a restricted data set that included only patients with complete data for all model variables (complete case analysis). We performed statistical analyses in R, version 3.1.0 (Vienna, Austria). All tests were two-sided, and we defined statistical significance by using $\alpha = .05$.

RESULTS

Patient Characteristics

A total of 10,125 patients with breast cancer were diagnosed and/or treated from 2005 to 2011 at either the University or Community health care system. Of these 10,125 patients, 2,776 had stage I to II HR-positive breast cancer; of these 2,776 patients, 2,418 (87.1%) were initially diagnosed with breast cancer in the participating health care systems and thus were classified as RS-eligible for our analysis. We defined as RS-eligible only patients diagnosed in the participating health care systems (thus excluding 358 patients who were diagnosed elsewhere but were later treated in the participating systems) because RS is indicated during the initial diagnostic period. Among this RS-eligible cohort, most patients (74.1%) were of non-Hispanic white race/ethnicity, followed by Asian (18.5%); most (59.3%) were of highest-quintile SES. A total of 773 RS-eligible patients (32.0%) were treated at the University system only, 1,367 (56.5%) were treated at the Community system only, and 278 (11.5%) were treated at both institutions (Table 1).

Of 2,418 RS-eligible patients, 377 (15.6%) received RS. Use of RS generally followed NCCN guidelines; among RS-eligible patients with no missing values for any variable, 92.0% had negative lymph nodes and 71.4% had stage I disease.⁵ The median RS among all RS-receipt patients was 17 (score range, 0 to 81; interquartile range, 13 to 22). The median RSs of all University, Community, and dual health care system RS-receipt patients were similar (17 [range, 0 to 55] v 17.5 [range, 0 to 81] v 16 [range, 0 to 56], respectively; Table 1). Among the 2,776 patients with stage I to II HR-positive disease, 2,408 (86.6%) had no missing variables; most missing variables were in the categories of grade and SES.

Multivariable Analysis

RS use among RS-eligible patients

On multivariable analysis, use of RS increased over time from 8.0% in 2005 to 2007 to 18.3% in 2008 to 2009 to 23.1% in 2010 to 2011 ($P < .001$; 2010-2011 v 2005-2007; odds ratio [OR], 3.68 [95% CI, 2.72 to 4.98]; Appendix Fig A1, online only). Clinicopathologic factors associated with lower RS use were grade ($P = .007$; grade 2 v 1: OR, 1.21 [95% CI, 0.92 to 1.59]; grade 3 v 1: OR, 0.67 [95% CI, 0.44 to 1.01]), involved lymph nodes ($P < .001$; 1 to 2 v none: OR, 0.19 [95% CI, 0.12 to 0.29]), and age ($P < .001$; < 40 v 50 to 64 years: OR, 0.42 [95% CI, 0.23 to 0.74]; ≥ 65 v 50 to 64 years: OR, 0.27 [95% CI, 0.17 to

0.43]). Race ($P = .62$) and SES ($P = .51$) were not associated with RS use. Patients who used both health care systems received the most RS testing ($P = .02$; v University: 1.73 [95% CI, 1.18 to 2.55]; Table 2).

Chemotherapy use among RS-eligible patients

Among RS-eligible patients, variables associated with receipt of chemotherapy included (monotonically) age ($P < .001$); age < 40 and 40 to 49 years (OR, 4.69 [95% CI, 2.89 to 7.62]) v 50 to 64 years (OR, 2.83 [95% CI, 2.14 to 3.73]), stage ($P < .001$); stage II v stage I: OR 3.62 [95% CI, 2.73 to 4.79], involved lymph nodes ($P < .001$; one v none: OR, 4.32 [95% CI, 3.12 to 5.97]), grade ($P < .001$; grade 2 v grade 1: OR, 3.12 [95% CI, 2.37 to 4.11]; grade 3 v grade 1: OR, 10.85 [95% CI, 7.44 to 15.82]), and race ($P = .03$; non-Hispanic white v non-Hispanic black: OR, 3.08 [95% CI, 1.26 to 7.55]; Table 2). Chemotherapy use decreased over time from 43.9% in 2005 to 2007 to 41.9% in 2008 to 2009 to 39.1% in 2010 to 2011, respectively ($P = .04$; Appendix Fig A1).

Chemotherapy use among RS-receipt patients

Among RS-receipt patients, higher RS was strongly associated with use of chemotherapy ($P < .001$; intermediate v low: OR, 3.66 [95% CI, 1.94 to 6.91]; high v low: OR, 49.39 [95% CI, 14.06 to 173.46]). Other correlates of chemotherapy use were similar to those found in the entire RS-eligible cohort. Involved lymph nodes were not associated with chemotherapy use among RS-receipt patients ($P = 0.21$; one to two nodes v none: OR, 1.96 [95% CI, 0.68 to 5.61]).

DISCUSSION

We developed an integrated EMR, SEER registry, and genomic testing database to investigate gene expression profiling for breast tumors and subsequent chemotherapy use in two health care systems that served the same geographic region in Northern California. Although one fourth of analyzed patients met guideline criteria for RS, we had evidence that only 16% of these RS-eligible patients actually received RS (compared with 20% in a recent NCCN study).³ Notably, use of RS increased over time (from 8.0% in 2005 to 2007 to 23.1% in 2010 to 2011), consistent with a recent study by Dinan et al² and likely reflecting providers' growing familiarity with the test. We identified a subset of patients for whom RS was used more often: patients treated in more than one health care system. These dual use patients differed from the larger group primarily by their higher SES and not by any clinically relevant

Table 1. Patient Characteristics

Characteristic	All Patients (N = 2,776)			RS-Eligible Patients (n = 2,418)*			RS-Receipt Patients (n = 377)		
	No.	%	Score Range	No.	%	Score Range	No.	%	Score Range
RS									
Not received	2,332	84.0		2,030	84.0		—	—	
Received	444	16.0		388	16.0				
Chemotherapy									
Not received	1,630	58.7		1,410	58.3		254	67.4	
Received	1,146	41.3		1,008	41.7		123	32.6	
RS (categorical)									
Low	229	8.2		200	8.3		194	52.0	
Intermediate	174	6.3		148	6.1		144	38.2	
High	41	1.5		40	1.7		37	9.8	
RS (continuous)†									
University		17.0	13.0-22.3	17.0	13.0-23.0	0-81	17.0	13.0-23.0	0-81
Community		17.0	13-23	17.0	13.3-23	0-55	17.0	13.3-23	0-55
Dual		17.5	13-23	17.0	13-23	0-81	17.0	13-24	0-81
		16.0	13-22	16.0	13-22.8	0-56	16.0	13-22	0-56
Age at diagnosis, years									
< 40	227	8.2		180	7.4		15	4.0	
40-49	659	23.7		566	23.4		122	32.4	
50-64	1,046	37.7		914	37.8		185	49.1	
≥ 65	844	30.4		758	31.3		55	14.6	
Year of diagnosis									
2005-2007	1,066	38.4		914	37.8		71	18.8	
2008-2009	774	27.9		663	27.4		120	31.8	
2010-2011	936	33.7		841	34.8		186	49.3	
Race									
Hispanic	155	5.6		144	6.0		20	5.3	
Non-Hispanic white	2,057	74.1		1,791	74.1		280	74.3	
Non-Hispanic black	41	1.5		33	1.4		3	0.8	
Non-Hispanic Asian/Pacific Islander	517	18.6		447	18.5		73	19.4	
Missing	6	0.2		3	0.1		1	0.3	

(continued on following page)

Table 1. Patient Characteristics (continued)

Characteristic	All Patients (N = 2,776)				RS-Eligible Patients (n = 2,418)*				RS-Receipt Patients (n = 377)			
	No.	%	Median	Score Range	No.	%	Median	Score Range	No.	%	Median	Score Range
Neighborhood SES statewide quintile												
1	60	2.2			54	2.2			10	2.7		
2	142	5.1			118	4.9			16	4.2		
3	295	10.6			255	10.5			34	9.0		
4	514	18.5			445	18.4			60	15.9		
5	1,624	58.5			1,435	59.3			236	62.6		
Missing	141	5.1			111	4.6			21	5.6		
Payer												
Private	1,927	69.4			1,655	68.4			306	81.2		
Medicare	658	23.7			586	24.2			44	11.7		
Military	12	0.4			11	0.5			1	0.3		
Uninsured/self-pay	12	0.4			11	0.5			0	0.0		
Public/Medicaid	96	3.5			89	3.7			11	2.9		
Missing	71	2.6			66	2.7			15	4.0		
Diagnosing health care system												
University	984	35.4			800	33.1			118	31.3		
Community	1,562	56.3			1,439	59.5			222	58.9		
Dual	190	6.8			179	7.4			37	9.8		
Treating health care system												
University	873	31.4			857	35.4			119	31.6		
Community	1,439	51.8			1,437	59.4			228	60.5		
Dual	128	4.6			124	5.1			30	8.0		
Health care system (combined)												
University	973	35.1			773	32.0			106	28.1		
Community	1,492	53.7			1,367	56.5			209	55.4		
Dual	293	10.6			278	11.5			62	16.4		
Stage												
I	1,590	57.3			1,378	57.0			269	71.4		
II	1,186	42.7			1,040	43.0			108	28.6		
No. of involved lymph nodes												
0	2,044	73.6			1,775	73.4			347	92.0		
1 or more	715	25.8			629	26.0			29	7.7		
Missing	17	0.6			14	0.6			1	0.3		

(continued on following page)

Table 1. Patient Characteristics (continued)

Characteristic	All Patients (N = 2,776)				RS-Eligible Patients (n = 2,418)*				RS-Receipt Patients (n = 377)			
	No.	%	Median	Score Range	No.	%	Median	Score Range	No.	%	Median	Score Range
Grade												
1	787	28.4		78.4	696	28.4		78.4	122	32.4		78.4
2	1,361	49.0		49.0	1,185	49.0		49.0	206	54.6		49.0
3	495	17.8		17.8	425	17.6		17.6	41	10.6		17.6
Missing	133	4.8		4.8	112	4.6		4.6	8	2.1		4.6
Histology												
Ductal	2,121	76.4		76.4	1,854	76.7		76.7	279	74.0		76.4
Lobular	291	10.5		10.5	251	10.4		10.4	51	13.5		10.5
Other	364	13.1		13.1	313	12.9		12.9	47	12.5		13.1
Laterality												
Unilateral	2,693	97.0		97.0	2,342	96.9		96.9	368	97.6		97.0
Bilateral	83	3.0		3.0	76	3.1		3.1	9	2.4		3.0

NOTE: Categorical variables are presented as numbers and percentages. Continuous variables are presented as medians and 25th and 75th percentiles (Q1-Q3).

Abbreviations: RS, recurrence score; SES, socioeconomic status.

*The RS-eligible cohort excludes patients who were diagnosed or treated outside the University or Community hospital system.

†RS summary statistics are calculated among RS recipients.

Table 2. Multivariable Logistic Regression Analyses With Multiple Imputation

Variable	RS-Eligible Patients (n = 2,418)						RS-Receipt Patients (n = 377)		
	Received RS			Chemotherapy			Chemotherapy		
	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P
Health care system			.02			.38			.11
University	Ref		Ref	Ref		Ref	Ref		Ref
Community	1.14	0.86 to 1.51	.36	0.86	0.66 to 1.11	.26	1.19	0.59 to 2.42	.62
Dual	1.73	1.18 to 2.55	.01	1.05	0.71 to 1.54	.82	2.47	1.01 to 6.04	.05
RS receipt						.11			
Not received	—		—	Ref		Ref	—		—
Received				0.78	0.58 to 1.05	.11			
RS									< .001
Low							Ref		Ref
Intermediate	—		—	—		—	3.66	1.94 to 6.91	< .001
High							49.39	14.06 to 173.46	< .001
Age at diagnosis, years			< .001			< .001			< .001
< 40	0.42	0.23 to 0.74	.0031	4.69	2.89 to 7.62	< .001	13.01	2.78 to 60.77	.001
40-49	1.13	0.85 to 1.49	.41	2.83	2.14 to 3.73	< .001	5.70	2.87 to 11.30	< .001
50-64	Ref		Ref	Ref		Ref	Ref		Ref
≥ 65	0.27	0.17 to 0.43	< .001	0.19	0.13 to 0.29	< .001	0.18	0.03 to 1.08	.06
Year of diagnosis			< .001			.12			.03
2005-2007	Ref		Ref	Ref		Ref	Ref		Ref
2008-2009	2.76	2.00 to 3.82	< .001	0.80	0.60 to 1.06	.12	2.10	0.89 to 4.98	.09
2010-2011	3.68	2.72 to 4.98	< .001	0.77	0.59 to 1.00	.05	0.86	0.38 to 1.95	.71
Race			.62			.03			.5
Non-Hispanic white	Ref		Ref	Ref		Ref	Ref		Ref
Hispanic	0.94	0.55 to 1.62	.83	1.30	0.79 to 2.13	.3	0.63	0.15 to 2.61	.52
Non-Hispanic black	0.44	0.12 to 1.53	.19	3.08	1.26 to 7.55	.01	1.90	0.11 to 32.21	.66
Non-Hispanic Asian/Pacific Islander	0.96	0.70 to 1.31	.79	1.29	0.96 to 1.72	.09	0.60	0.28 to 1.27	.18
Neighborhood SES statewide quintile			.51			.41			.57
1	Ref		Ref	Ref		Ref	Ref		Ref
2	0.81	0.32 to 2.04	.65	1.79	0.72 to 4.47	.21	1.76	0.17 to 18.00	.63
3	0.72	0.32 to 1.64	.43	1.12	0.49 to 2.57	.79	0.87	0.11 to 7.04	.89
4	0.67	0.30 to 1.47	.31	1.30	0.58 to 2.95	.52	2.24	0.30 to 16.92	.43
5	0.87	0.40 to 1.89	.73	1.11	0.51 to 2.44	.79	1.24	0.19 to 8.06	.82
Insurance			.4			.19			.17
Private	Ref		Ref	Ref		Ref	Ref		Ref
Medicare	0.85	0.51 to 1.40	.51	0.67	0.44 to 1.03	.07	3.30	0.49 to 22.26	.22
Other	0.66	0.34 to 1.27	.21	0.87	0.51 to 1.48	.6	3.22	0.71 to 14.56	.13
Stage			.42			< .001			.01
I	Ref		Ref	Ref		Ref	Ref		Ref
II	0.88	0.66 to 1.19	.42	3.62	2.73 to 4.79	< .001	2.74	1.34 to 5.62	.01
No. of involved lymph nodes			< .001			< .001			.21
0	Ref		Ref	Ref		Ref	Ref		Ref
1-2	0.19	0.12 to 0.29	< .001	4.32	3.12 to 5.97	< .001	1.96	0.68 to 5.61	.21

(continued on following page)

Table 2. Multivariable Logistic Regression Analyses With Multiple Imputation (continued)

Variable	RS-Eligible Patients (n = 2,418)						RS-Receipt Patients (n = 377)		
	Received RS			Chemotherapy			Chemotherapy		
	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P
Grade			.007			< .001			.01
1	Ref		Ref	Ref		Ref	Ref		Ref
2	1.21	0.92 to 1.59	.17	3.12	2.37 to 4.11	< .001	2.70	1.33 to 5.51	.01
3	0.67	0.44 to 1.01	.06	10.85	7.44 to 15.82	< .001	4.53	1.58 to 12.94	.005
Histology			.22			.06			.02
Ductal	Ref		Ref	Ref		Ref	Ref		Ref
Lobular	1.39	0.96 to 2.02	.08	0.75	0.51 to 1.11	.15	0.34	0.11 to 0.98	.05
Other	1.06	0.75 to 1.52	.73	0.69	0.49 to 0.98	.04	0.35	0.13 to 0.90	.03
Laterality			.79			.52			
Unilateral	Ref		Ref	Ref		Ref	—*		
Bilateral	1.09	0.56 to 2.13	.79	0.81	0.42 to 1.55	.52			

NOTE. Logistic regression models were fit to multiply-imputed data by using five imputations. Coefficient *P* values and variable *P* values (in bold) were calculated by using Wald's *z* test.

Abbreviations: OR, odds ratio; Ref, reference; RS, recurrence score; SES, socioeconomic status.

*Categories were collapsed because there were few observations.

characteristics such as cancer prognostic factors or median RS (16 among dual users *v* 17 among all patients). Reassuringly, however, RS-receipt patients had no differential use of chemotherapy associated with sociodemographic characteristics that were not pertinent to their cancer care (eg, race/ethnicity, SES, or health care system). We therefore speculate that the emerging genomic technologies used to tailor medical therapies might serve to optimize cancer treatment across health care settings.

Several prior studies have evaluated RS use and subsequent patterns of care.^{2,3,12-14,27} A novel aspect of our study was our incorporation of patient-level numerical RS values made available through a partnership with investigators at the testing laboratory (GHI). Having these RS values enabled a quantitative comparison of tumor biology across patient subgroups and care settings. Accordingly, our finding that chemotherapy use followed RS practice guidelines across institutions evinces an equally high quality of care within the Community and University settings we studied. Previous studies found that RS testing alters treatment recommendations for 20% to 40% of eligible patients with breast cancer, with an associated reduction in chemotherapy use over time since the widespread adoption of RS in clinical practice.^{2,4,8,12-14,28-30} Our study echoes these findings; specifically, RS-receipt patients had a trend toward lower odds of

receiving chemotherapy, although this finding did not reach statistical significance. We previously reported on breast cancer care in these Community and University health care settings and found that patients who used the dual health care system had significantly greater use of every studied intervention, including bilateral mastectomy, magnetic resonance imaging scans, positron emission tomography, and *BRCA1/2* genetic testing.¹⁵ After integrating genomic data, we now add RS to the list of more frequently accessed interventions. Importantly, however, we observed that chemotherapy use followed guidelines among all RS-receipt patients, regardless of health care setting, race/ethnicity, or SES. These results echo recent studies on the use of RS,^{2,3,27,31} and suggest that if RS is performed, providers generally follow its recommendations. Tailored genomic testing thus has the potential to reduce disparities in cancer treatment, including the possible overuse of chemotherapy.

Limitations of our study include restrictions on available EMR data. Although structured data from billing, drug ordering, and administration records are readily extracted, the nuances of clinical decision making are often buried in dictated notes, and thus are challenging to extract. The Oncoshare database draws from a single catchment area in Northern California, focusing on two health care systems that disproportionately represent non-Hispanic whites and Asians with

high SES, and thus results may not be broadly generalizable. Given the patient demographics of this study, there may be insufficient statistical power to detect an association between RS receipt and chemotherapy use in less represented patients (eg, those of nonwhite, non-Asian race/ethnicity who have lower SES). The Oncoshare database currently lacks patient-reported information about care decisions and cannot capture use of additional health care systems outside the participating University and Community systems. As with all observational studies, causality inferred from associations can be misleading as a result of uncontrolled confounding or selection biases. For example, RS was received by only a fraction of eligible patients (15.6%); there may be unmeasured clinicopathologic risk factors that differ between this select group and the general patient population. Reassuringly, we did not find any evidence of a difference in risk profile between the patients who received the RS versus those who did not, yet there may be unmeasured prognostic differences. Accordingly, we are now investigating factors that trigger the health care behavior of patients accessing services in both University and Community health care systems. We lacked data on at least two variables for 13% of patients (primarily SES and histologic grade); notably, however, complete case analysis and MI of missing data were comparable with regard to factors associated with RS and chemotherapy use. Because the validity of findings from a complete case analysis rely on stringent assumptions regarding missing data (eg, that those with complete data are similar to those with incomplete data), and validity of findings from an MI-based analyses rely on more flexible assumptions about missing data (eg, that missing data are related to observed characteristics only), we prefer to interpret findings on the basis of the MI-based models.

In this real-world study of more than 2,000 patients with breast cancer treated in Community and University health care settings, RS use was low (15.6%) but increased over time, which likely reflects a gradual dissemination of genomic testing into routine care. There was greater use of RS among patients who accessed more than one health care setting, yet no difference in the use of RS after chemotherapy among patients treated in Community, University, or both settings. Collaboration between academic, community, and industry investigators enabled this novel linkage of EMR, SEER registry, and gene expression data. Our prior work using this integrated approach consistently identified a cluster of apparent overuse of diagnostic and treatment interventions by patients who accessed health care across health care systems¹⁵; here, we tested the

hypothesis that such usage patterns would extend to the receipt of chemotherapy after genomic profiling. However, our results suggest instead that ongoing efforts toward precision medicine, such as use of the RS, may reduce unwarranted variation and disparities in cancer care. This may be a potential mechanism by which patients who otherwise seek more resources can have their health care use mitigated by reassuring results of objective genomic testing. **JOP**

Acknowledgment

Supported by the Susan and Richard Levy Gift Fund; the Suzanne Pride Bryan Fund for Breast Cancer Research; the Breast Cancer Research Foundation; Grants No. 16OB-0149 and 19IB-0124 from the Regents of the University of California's California Breast Cancer Research Program; the Stanford University Developmental Research Fund; Contract No. HHSN261201000140C from the National Cancer Institute's (NIH's) Surveillance, Epidemiology, and End Results (SEER) program (Cancer Prevention Institute of California); and by Clinical and Translational Science Award No. UL1 RR025744a from the NIH. The collection of cancer incidence data used in this study was supported by the California Department of Health Services as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885, by Contract No. HHSN261201000140C from NIH's SEER program, Contract No. HHSN261201000035C (University of Southern California), Contract No. HHSN261201000034C (Public Health Institute), and Agreement No. 1U58 DP000807-01 from the Centers for Disease Control and Prevention's National Program of Cancer Registries (Public Health Institute).

Presented at the 50th Annual Meeting of the American Society of Clinical Oncology, Chicago, IL, May 30-June 3, 2014.

A.A. and M.M. contributed equally to this work.

The ideas and opinions expressed herein are those of the authors; endorsement by the University of California, State of California, the California Department of Health Services, the National Cancer Institute, or the Centers for Disease Control and Prevention, or their contractors and subcontractors is not intended nor should it be inferred.

Authors' Disclosures of Potential Conflicts of Interest

Disclosures provided by the authors are available with this article at jop.ascopubs.org.

Author Contributions

Conception and design: Anosheh Afghahi, Caroline A. Thompson, Manisha Desai, Peter P. Yu, Amar K. Das, Harold S. Luft, Amy P. Sing, Allison W. Kurian

Financial support: Allison W. Kurian

Provision of study materials or patients: Scarlett L. Gomez

Collection and assembly of data: Maya Mathur, Caroline A. Thompson, Aya Mitani, Manisha Desai, Tina Seto, Cliff Olson, Pragati Kenkare, Scarlett L. Gomez, Amar K. Das, Amy P. Sing, Allison W. Kurian

Data analysis and interpretation: Anosheh Afghahi, Maya Mathur, Caroline A. Thompson, Aya Mitani, Joseph Rigdon, Manisha Desai, Peter P. Yu, Monique A. de Bruin, Tina Seto, Scarlett L. Gomez, Amar K. Das, Harold S. Luft, George W. Sledge, Jr, Amy P. Sing, Allison W. Kurian

Manuscript writing: All authors

Final approval of manuscript: All authors

Corresponding author: Allison W. Kurian, MD, MSc, Stanford University School of Medicine, HRP Redwood Building, Rm T254A, 150 Governor's Lane, Stanford, CA 94305-5405; e-mail: akurian@stanford.edu.

References

1. Siegel R, Naishadham D, Jemal A: Cancer statistics, 2013. *CA Cancer J Clin* 63: 11-30, 2013

2. Dinan MA, Mi X, Reed SD, et al: Initial trends in the use of the 21-Gene Recurrence Score Assay for patients with breast cancer in the Medicare population, 2005-2009. *JAMA Oncol* 1:158-166, 2015
3. Hassett MJ, Silver SM, Hughes ME, et al: Adoption of gene expression profile testing and association with use of chemotherapy among women with breast cancer. *J Clin Oncol* 30:2218-2226, 2012
4. Ademuyiwa FO, Miller A, O'Connor T, et al: The effects of Oncotype DX recurrence scores on chemotherapy utilization in a multi-institutional breast cancer cohort. *Breast Cancer Res Treat* 126:797-802, 2011
5. Carlson RW, Allred DC, Anderson BO, et al: Invasive breast cancer. *J Natl Compr Canc Netw* 9:136-222, 2011
6. Gianni L, Zambetti M, Clark K, et al: Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer. *J Clin Oncol* 23:7265-7277, 2005
7. Habel LA, Shak S, Jacobs MK, et al: A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients. *Breast Cancer Res* 8:R25, 2006
8. Hornberger J, Cosler LE, Lyman GH: Economic analysis of targeting chemotherapy using a 21-gene RT-PCR assay in lymph-node-negative, estrogen-receptor-positive, early-stage breast cancer. *Am J Manag Care* 11:313-324, 2005
9. Paik S, Shak S, Tang G, et al: A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 351:2817-2826, 2004
10. Paik S, Tang G, Shak S, et al: Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* 24:3726-3734, 2006
11. Sparano JA, Gray RJ, Makower DF, et al: Prospective validation of a 21-gene expression assay in breast cancer. *N Engl J Med* 373:2005-2014, 2015
12. Shak S, Petkov VI, Miller DP, et al: Breast cancer specific survival in 38,568 patients with node negative hormone receptor positive invasive breast cancer and Oncotype DX recurrence score results in the SEER database. 38th Annual San Antonio Breast Cancer Symposium, San Antonio, TX, December 8-12, 2015 (abstr P5-15-01)
13. Stemmer SM, Steiner M, Rizel S, et al: Real-life analysis evaluating 1594 N0/Nmic breast cancer patients for whom treatment decisions incorporated the 21-gene recurrence score result: 5-year KM estimate for breast cancer specific survival with recurrence score results ≤ 30 is $>98\%$. 38th Annual San Antonio Breast Cancer Symposium, San Antonio, TX, December 8-12, 2015 (abstr P5-08-02)
14. Swain SM, Nunes R, Yoshizawa C, et al: Quantitative gene expression by recurrence score in ER-positive breast cancer, by age. *Adv Ther* 32:1222-1236, 2015
15. Kurian AW, Mitani A, Desai M, et al: Breast cancer treatment across health care systems: Linking electronic medical records and state registry data to enable outcomes research. *Cancer* 120:103-111, 2014
16. Weber SC, Seto T, Olson C, et al: Oncoshare: Lessons learned from building an integrated multi-institutional database for comparative effectiveness research. *AMIA Annu Symp Proc* 2012:970-978, 2012
17. Thompson CA, Kurian AW, Luft HS: Linking electronic health records to better understand breast cancer patient pathways within and between two health systems. *EGEMS (Washington, DC)* 3:1127, 2015
18. Lowe HJ, Ferris TA, Hernandez PM, et al: STRIDE: An integrated standards-based translational research informatics platform. *AMIA Annu Symp Proc* 2009:391-395, 2009
19. Yost K, Perkins C, Cohen R, et al: Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. *Cancer Causes Control* 12:703-711, 2001
20. National Comprehensive Cancer Network (NCCN): NCCN Clinical Practice Guidelines in Oncology: Breast Cancer, Version 1, 2016. <http://www.nccn.org>
21. Albain KS, Barlow WE, Shak S, et al: Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: A retrospective analysis of a randomised trial. *Lancet Oncol* 11:55-65, 2010
22. van Buuren S: Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res* 16:219-242, 2007
23. van Buuren S. *Flexible Imputation of Missing Data*. Boca Raton, FL, CRC Press, 2012
24. Barnard J, Rubin DB: Small-sample degrees of freedom with multiple imputation. *Biometrika* 86:948-955, 1999
25. Mitani A, Kurian AW, Das AK, et al: Navigating choices when applying multiple imputation in the presence of multi-level categorical interaction effects. *Stat Methodol* 27:82-99, 2015
26. van Buuren S, Groothuis-Oudshoorn K: mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 45:1-67, 2011. <http://doc.utwente.nl/78938/1/Buuren11mice.pdf>
27. Potosky AL, O'Neill SC, Isaacs C, et al: Population-based study of the effect of gene expression profiling on adjuvant chemotherapy use in breast cancer patients under the age of 65 years. *Cancer* 121:4062-4070, 2015
28. Klang SH, Hammerman A, Liebermann N, et al: Economic implications of 21-gene breast cancer risk assay from the perspective of an Israeli-managed health-care organization. *Value Health* 13:381-387, 2010
29. Lo SS, Mumby PB, Norton J, et al: Prospective multicenter study of the impact of the 21-gene recurrence score assay on medical oncologist and patient adjuvant breast cancer treatment selection. *J Clin Oncol* 28:1671-1676, 2010
30. Lyman GH, Cosler LE, Kuderer NM, et al: Impact of a 21-gene RT-PCR assay on treatment decisions in early-stage breast cancer: An economic analysis based on prognostic and predictive validation studies. *Cancer* 109:1011-1018, 2007
31. Roberts MC, Weinberger M, Dusetzina SB, et al: Racial variation in adjuvant chemotherapy initiation among breast cancer patients receiving Oncotype DX testing. *Breast Cancer Res Treat* 153:191-200, 2015

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST**Use of Gene Expression Profiling and Chemotherapy in Early-Stage Breast Cancer: A Study of Linked Electronic Medical Records, Cancer Registry Data, and Genomic Data Across Two Health Care Systems**

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or jop.ascopubs.org/site/misc/ifc.xhtml.

Anosheh Afghahi

No relationship to disclose

Maya Mathur

No relationship to disclose

Caroline A. Thompson

No relationship to disclose

Aya Mitani

No relationship to disclose

Joseph Rigdon

No relationship to disclose

Manisha Desai

No relationship to disclose

Peter P. Yu

Stock or Other Ownership: ContraFect, Citrix Systems, EMC, Google, IBM, Oracle, FireEye, Apple

Research Funding: Berg Pharma (Inst)

Monique A. de Bruin

Consulting or Advisory Role: Varian Medical Systems

Tina Seto

No relationship to disclose

Cliff Olson

No relationship to disclose

Pragati Kenkare

No relationship to disclose

Scarlett L. Gomez

Employment: Eurofins (I), BioInspire Technologies (I)

Stock or Other Ownership: Amgen (I)

Research Funding: Genentech (Inst)

Travel, Accommodations, Expenses: Genentech

Amar K. Das

No relationship to disclose

Harold S. Luft

No relationship to disclose

George W. Sledge, Jr

Leadership: Syndax Pharmaceuticals

Stock or Other Ownership: Syndax Pharmaceuticals

Honoraria: Symphogen

Consulting or Advisory Role: Symphogen, Nektar Therapeutics, Radius

Research Funding: Genentech (Inst)

Travel, Accommodations, Expenses: GlaxoSmithKline, Nektar Therapeutics

Amy P. Sing

Employment: Genomic Health

Stock or Other Ownership: Genomic Health

Allison W. Kurian

Research Funding: Myriad Genetics (Inst), Invitae (Inst), Ambry Genetics (Inst), GeneDx (Inst), Genomic Health (Inst)

Appendix

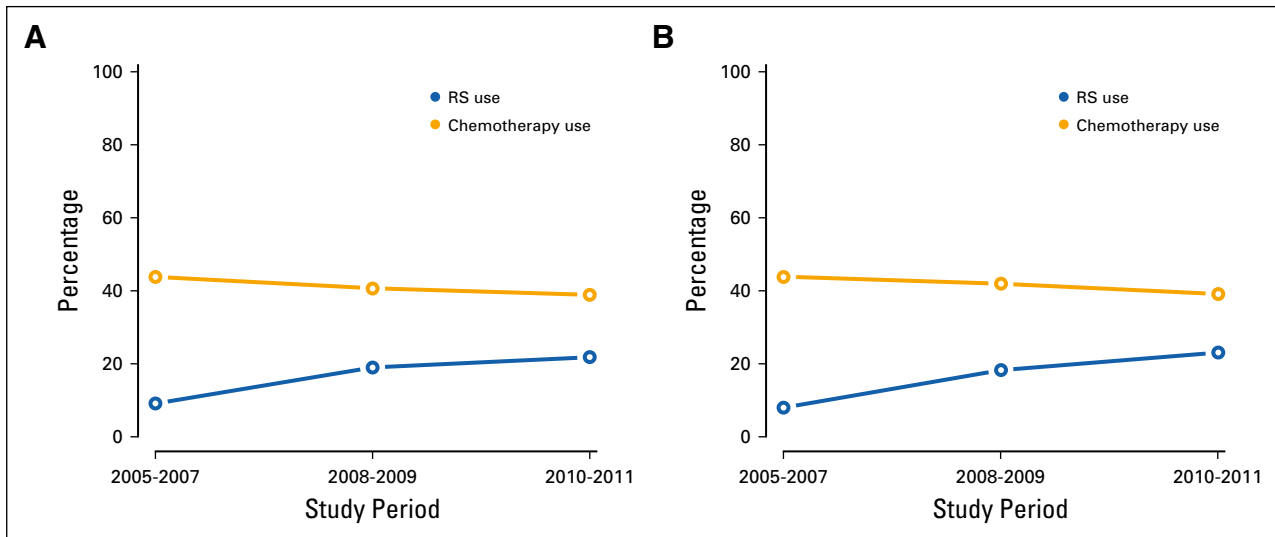


FIG A1. Recurrence score (RS) and chemotherapy use by study period in (A) the overall study population and (B) RS-eligible study participants.