ISSN: 2755-984X

Journal of Tumor Science Research

SCIENTIFIC Research and Community

Research Article

Open Access

Overtreatment in Hormone Receptor Positive, HER2-, Axillary Node Negative, Non-Metastatic Early Breast Cancer

Samuel Schick[1]* and Francesco Schettini[2]

[1]School of Medicine, West Virginia Univerisity, Morgantown WV, USA

[2]Deptartment of Oncology, Hospital Clinic Barcelona, Spain

ABSTRACT

HR+ HER2- T1 N0 M0 breast cancer has very low 5-year postoperative risk of relapse (<5%) with current SoC treat- ment regimens: endocrine monotherapy or chemoendocrine com- bination therapy. Minimally aggressive treatment is ideal. Absent gene-profiling, histopathological features of the tumor are examined to predict risk of relapse. This paper establishes statistical evidence for use of an empirical prognostic index as a risk-based clinical decision aid where gene profiles are unavailable. Methods: A retrospective cohort (n=965) observed for 5 years was propensity score (PS) matched to correct baseline risk differences between chemoendocrine therapy and endocrine monotherapy treatment groups. PS were generated from logistic regression of pertinent histopathological covariates available to a clinician at the time of diagnosis onto treatment designation. Comparator groups were fur- ther equilibrated using Synthetic Minority Over-Sampling (SMOTE). Estimated treatment effects were assessed after propensity score- matching (PSM) and after SMOTE-normalized re-sampling. Results: Propensity score-matched groups (n=262) showed no difference in relapse rates between groups. Significant differences in histopatho- logical covariates remained between groups after PSM. SMOTE further minimized inter-group differences. Corresponding propen- sity score quartiles of the SMOTE groups (n=1000) were tested for proportional differences in relapse. No proportional differences in relapses were observed between groups in Q1-Q3. Proportionally, significantly more relapses occurred in the Q4 endocrine therapy group than in the Q4 chemoendocrine group. Conclusion: After matching, the PSM sample failed to produce generalizable results due to significant group imbalance for multiple predictor variables, whereas SMOTE re-sampling provided adequately risk-matched groups (standardized differences<0.2) for treatment effect anal- ysis. Individuals with HR+ HER2- T1 N0 M0 breast cancer are over-treated if they have PS<0.75, and are prescribed chemoen- docrine therapy, as no additional benefit was observed from this more aggressive treatment regimen. Conversely, individuals are under-treated, and at a significantly higher risk of relapse, if they have PS>0.75 and are not prescribed chemotherapy.

***Corresponding author**

Samuel Schick MD, School of Medicine, West Virginia Univerisity, Morgantown WV, USA. E-mail: ss0226@wvu.edu

Received: September 04, 2022; **Accepted:** September 12, 2022; **Published:** September 20, 2022

Keywords: Propensity Score Matching, SMOTE, Survival Analysis, HR+, HER2-, Breast Cancer, Treatment De-Escalation

Abbreviations

SoC: Standard of Care
HR+: Hormone Receptor Positive
HER2-: Human Epidermal Growth Factor Receptor 2 Negative
T1: Tumor Dimensions <2cm
N0: Axillary Node Negative
M0: Non-Metastatic Abbreviations
Age: Age at Diagnosis (Years)
G: Tumor Grade
ER: Percent of Staged Cells Expressing Estrogen Receptor Expression
PgR: Percent of Staged Cells Expressing Progesterone Receptor Expression
Ki67: Percent of Staged Cells Expressing MKI67
LocalR: Individuals with Observed Local Metastasis or Tumor Progression
DistantR: Individuals with Distal Observed Metastasis or Tumor Progression
Relapse: Individuals with Observed Local and/or Distal Metastasis or Tumor Progression
obsT: Observation Time (days).

Introduction

Estrogen and/or progesterone hormone receptor positive (HR+), HER2 oncogene overexpression negative (HER2-), axillary lymph node negative, (N0) non-metastatic (M0), T1 (tumors size <2cm) breast cancer is a subtype of hormone- sensitive breast cancer. For reasons discussed below, patients with this subtype of breast cancer are at risk for over- treatment. Endocrine (hormone) therapies are viable treatment options for hormone responsive tumors [1]. Agnostic of treatment method, either endocrine or chemoendocrine, 5-year disease-free survival (DFS) of HR+ HER2- T1 N0 M0 breast cancer following surgery is over 90%; chemotherapy is not generally recommended as a first line of treatment by physicians if a gene profile test is available to rule out risk of endocrine resistance, a common biochemical change that can lead to relapse during or after endocrine therapy [2, 3]. Hence, there exists a clinical need for accurate prediction of relapse for the purposes of reducing the detriments associated with both over-treatment and under-treatment.

National Comprehensive Cancer Network (NCCN) guidelines and current literature encourage adopting prognostic systems based on gene expression profiling to identify patients at major risk of relapse, defined as detectable breast tumor growth, in order to decide whether to administer adjuvant chemotherapy in addition to hormonal therapy [4, 5]. However, genetic profiling tests, such as Oncotype Dx, are not yet widespread outside of the United States. In the absence of gene profiling resources, the treatment decision to add chemotherapy is largely based on prognostic factors discernible through immunohistochemistry of tumor tissue samples. Due to the favorable prognosis of HR+ HER2- T1 N0 M0 breast cancer, weighed against the detrimental effects of chemotherapy, consensus on treatment strategies has shifted toward emphasizing the de-escalation of treatment [6].

In the absence of gene-profiling, accurate relapse risk prediction based on pathology and histology is essential for adequate and appropriate treatment of HR+, HER2-, T1, N0, M0 breast cancer. Based on the 5-year multi-center observational data available, this paper proposes and statistically tests an empirical prognostic scoring system to aid in the decision to prescribe or withhold from prescribing post-surgical chemotherapy, in addition to hormonal therapy, for treatment of HER-, T1, N0, M0, breast cancer.

Methods

Study Design

This multi-center retrospective cohort study includes 5 years of post-surgery observation of 965 patients, with HR+ HER2- T1 N0 M0 breast cancer, in Italy (University of Naples "Federico II", National Cancer Institute "G. Pascale", Naples, AORN "A. Cardarelli" Hospital, Naples, and the "Santa Maria della Misericordia" University Hospital, Udine). Variables include survival status, patient age, last date of observation, surgery date, metastasis (local/distal) and tumor growth status, tumor grading, HR expression (estrogen and progesterone), tumor location (apical/luminal), and ki67 proliferation scoring (0 for low proliferation index and 1 for high proliferation index). These correspond to the prognostic factors currently deemed of clinical significance by the NCCN guidelines and 2013 St. Gallen consensus on breast cancer: location (luminal or apical), histological grade (G), proliferation index (Ki67), levels of estrogen receptor (ER) and progesterone receptor (PgR) and T1 tumor dimension (between 0.1-0.5 cm, between 0.5-1cm, between 1-2 cm) [4, 6].

Outcome Measures

The endpoints analyzed in this study are disease free survival, relapse status (distal, local, and overall), during an average 5-year observation period immediately following first surgical intervention. Mortality was not included due to the rarity of the event: only 5 individuals died of any cause, related or unrelated to breast cancer, out of 965 individuals observed. Survival was not estimated due to sample size limitations. Subgroup analysis of relapse location (distal versus local) related to treatment allocation was performed but resulted insignificant findings, likely due to small event totals. Overall relapse, a binary variable indicating either local or distal relapse, was used as the main outcome of interest throughout statistical analysis. The primary predictor of interest will be propensity score. This propensity score is composed of age, progesterone receptor expression, estrogen receptor expression, Ki67 proliferation index, tumor dimensions and tumor grade.

Observation time was available for all subjects, but was not included in forming the predictor variable (propensity score). because treatment duration is not known to the clinician at the time of diagnosis. Predictor variables necessarily utilize known information available to the clinician to assess patient risk of relapse following surgery. All patients underwent hormone adjuvant therapy. While clinically relevant, the variable was not included as a parameter in the either regression model because it cannot contribute to distinguishing outcome probabilities between the two comparator groups.

Treatment allocation was not randomized. Balance in predictor variables between treatment arms is used as a proxy for validity since samples with differentially distributed have potential for confounding and masking of true treatment effect. Predictor variables with standardized differences between treatment groups of above 0.2 were considered compromising to the validity of the statistical analysis.

Statistical Analysis

The original endocrine (n=695) and chemoendocrine (n=270) groups were followed for a median time of 68.5 months and 85 months, respectively. During this time, 18 relapses were observed in each group, which equates to a 2.59% and 6.67% 5-year relapse rate in the endocrine and chemoendocrine groups, respectively. The median endocrine subject was 61 years old versus 50 years old for the chemoendocrine group. Ki67 proliferation index was double (20%) in the chemoendocrine group relative to the endocrine group (10%). Progesterone and estrogen receptor expression was 20% higher in the endocrine group, at 70% and 90% respectively, relative to the chemoendocrine group.

To mitigate significant imbalances the characteristics of the two treatment groups, the original sample was propensity score-matched then bootstrapped (sampled with replacement) using Synthetic Minority Oversampling (SMOTE).

Propensity score matching (PSM) is a statistical method of matching individuals between groups by ordering samples based on the log-likelihood of a binary outcome based on a combination of predictor variables. Prior to matching, the predictor variables are regressed onto the binary outcome variable. The propensity to have the attribute, in this instance a chemotherapy treatment regimen (CT=1), is matched between the two groups to assess a separate outcome measure. Data points were matched using a 'nearest neighbor' matching algorithm. Unmatched data points in both groups were discarded after 1:1 matching, resulting in 262 propensity-matched pairs between groups. The MatchIt package (version 3.0.2) in R was used for matching.

After PSM and SMOTE, treatment effects were tested with a t-test for proportions, and utilizing a Bernoulli correc- tion was applied to account for increased probability of committing a type I error when multiple samples were tested simultaneously. Prerequisite to PSM statistical inference is that the comparator groups must have roughly equivalent distributions of predictor covariates. While PSM is not an infallible substitute for a well-conducted randomized clinical trial, the statistical testing of well-balanced PSM samples can be used as inference to assess treatment effects in the absence of a causal based experiment.

Survival analysis with the Kaplan-Meier curves and a log-rank test was performed to test the effect of chemotherapy on relapse free survival in the propensity-matched sample [7-10]. Given that the propensity scores are constructed using a multivariate logistic regression, a univariate Kaplan-Meier regression on these scores roughly equates to a multivariate analyses with Cox regression in which the covariates are prognostic index, age and chemotherapy,

grading, Ki67, tumor dimension and hormone receptor levels.

Lastly, bootstrapping was performed on the propensity score matched sample, in order to better estimate the empiri- cal distribution of the control and treatment groups. Bootstrapping is a widely accepted re-sampling technique used to improve estimation and prediction capability by simulating larger sample sizes than are available to the researcher [11, 12]. The bootstrapping was conducted with an implementation of the SMOTE algorithm [13]. This technique was utilized to further enhance the estimation ability of bootstrapping by minimizing the residual effects of class imbalance on the propensity matched sample.

Results

Table 1: Original sample: predictor and response variables, grouped by treatment designation

	Endocrine Therapy	Chemothrapy + Endocrine Therapy	P-value	Std. Difference
n	695	270		
Age (mean (sd))	60.20 (10.64)	50.86 (8.97)	<0.001	0.949
G (mean (sd))	1.88 (0.59)	2.43 (0.57)	<0.001	0.961
ER (mean (sd))	82.66 (16.00)	63.75 (27.41)	<0.001	0.843
PgR (mean (sd))	60.78 (32.18)	49.54 (31.66)	<0.001	0.352
Ki67 (mean (sd))	13.34 (10.69)	23.25 (15.21)	<0.001	0.753
T1abc			<0.001	0.647
0	59 (8.5)	4 (1.5)		
1	269 (38.7)	46 (17.0)		
2	367 (52.8)	220 (81.5)		
Luminal (%)	270 (38.8)	180 (66.7)	<0.001	0.58
Relapse (%)	18 (2.6)	18 (6.7)	0.005	0.195
LocalR (%)	8 (1.2)	6 (2.2)	0.342	0.083
DistantR (%)	11 (1.6)	11 (4.1)	0.037	0.151
Death (%)	2 (0.3)	4 (1.5)	0.097	0.128
obsT (mean (sd))	2146.46 (893.67)	2613.54 (1044.86)	<0.001	0.48

Clear and highly significant differences exist between groups in age, tumor grade and dimension, hormone expres- sion, and proliferation index. This imbalance prevents from drawing conclusions about the outcome of interest, relapse free survival, due to the potential issue of confounding. Hence, propensity score matching is utilized to minimize group imbalance and simulate random group allocation.

Table 2: Generalized (binomial/logit) linear model: CT = Age + G + ER + PgR + Ki67 + T1abc + Luminal

Variable	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.4	0.73	1.9	0.052
Age	-0.1	0.011	-9.3	1.4e-20
G	1.4	0.18	8.1	6.4e-16
ER	-0.034	0.0049	-7	2.9e-12
PgR	-0.01	0.0037	-2.8	0.0052
Ki67	0.045	0.011	4.2	3.2e-05
T1abc	1.5	0.21	7.4	1.6e-13
Luminal	-0.046	0.28	-0.16	0.87

Logistic Regression summary: logistic regression performed on the dataset to estimate propensity to receive chemotherapy, in addition to hormonal therapy, following surgical removal of the tumor(s). The logistic regression used to generate propensity scores is an estimate of cumulative clinical decisions, risk of relapse quantification, made by physicians of individuals in the cohort. This regression shows potential for use as prognostic index for risk of relapse in HR+, HER2-, T1, N0, M0 breast cancer. Statistical analysis in this paper suggests a cutoff at PS=0.75, above which chemoendocrine therapy is beneficial for the patient.

Equation 1. The log-link function:

$$ln\left(\frac{e^{x}}{1-e^{x}}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_n x_n \quad (1)$$

The log-link function is the mathematical transformation of this regression into a propensity score from which treat- ment and control groups can be constructed. The score will range from 0 to 1 and represents, for a given combination of covariate levels, the log probability ratio that a doctor will treat a patient with chemotherapy

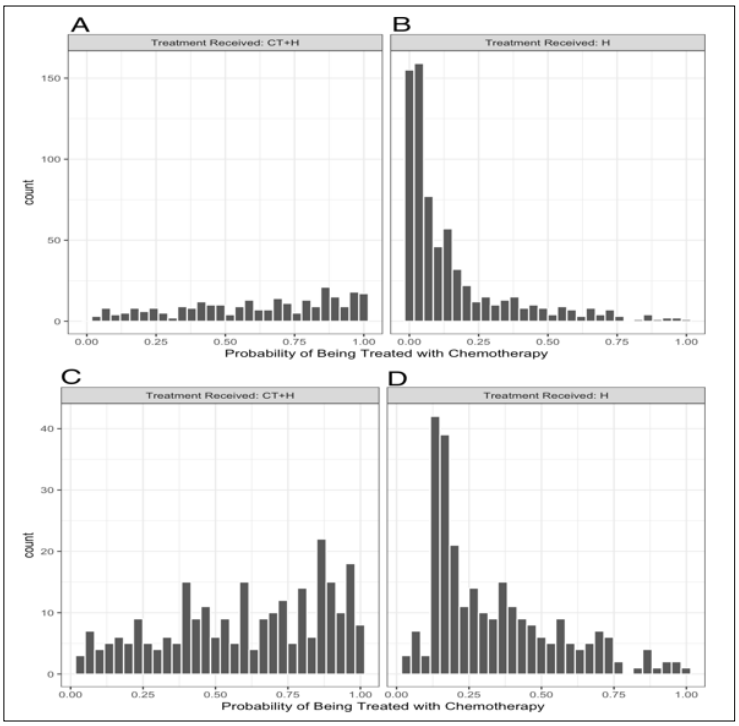

Figure 1: Histogram of Sample Propensity Scores Before A, B and After C, D Matching

As illustrated in Figure 1 A and B the original sample contained heavily imbalanced groups, which confirms the necessity of employing PSM techniques. The large class imbalance in the unmatched data translated to smaller residual imbalance in the matched sample, Figure 1 C and D, in the propensity range of between 0.1 and 0.2 most visibly.

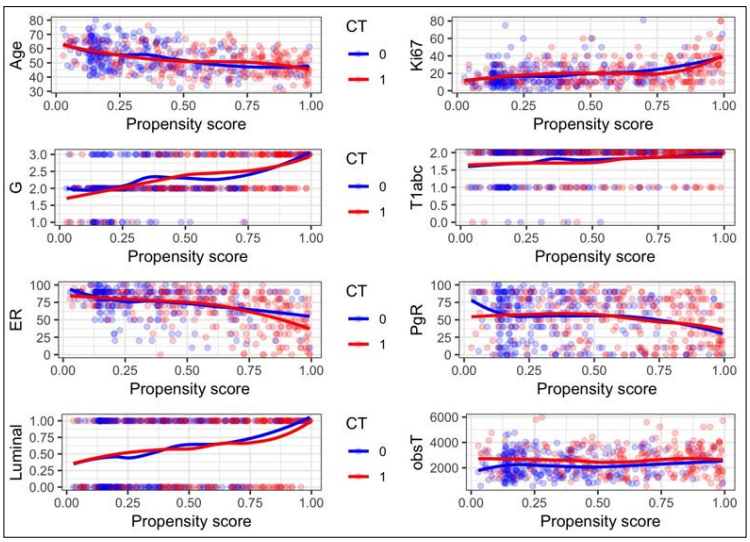

Figure 2: Propensity Distributions of Observed Covariates

Despite differences in the distribution of matched propensity scores in Figure 1, the locally-weighted average of each variable across the entire propensity score range remains similar between the hormonal therapy mono-treatment group (blue) and the chemotherapy with hormonal therapy group (red) in Figure 2. It should be noted that while the local averages of these variables are comparable, the frequency relative frequency distributions of severable variables change across the range of the propensity score. The relative frequency of chemotherapy treatment increases in ER, PgR, and Ki67, and decreases with Age, as propensity score increases. A two-tailed two sample t-test at the alpha = 0.05 level, after matching, the proportion of relapse is no longer significantly different between treatment and control groups (P-val = 0.3555). The original sample contained a significant difference in relapse rate across treatment groups, with the cancer patients experiencing significantly more relapses, proportionally, than the hormonal therapy monotherapy group as demonstrated by a two-tailed t-test in Table 3 (P-val=0.01316). Since the statistically significant difference vanished once the sample was matched, it is likely that the difference was an artifact resulting from uncontrolled class imbalance that has been corrected, to an extent, in the process of creating the propensity matched sample.

Table 3: Control and Treatment Group Means in Propensity-Matched Sample. Matched sample characteristics of interest, grouped by treatment arm

	Endocrine Therapy	Chemothrapy + Endocrine Therapy	P-value	Std. Difference
n	262	262		
Age (mean (sd))	54.56 (9.76)	51.19 (8.82)	<0.001	0.362
G (mean (sd))	2.14 (0.60)	2.42 (0.57)	<0.001	0.476
ER (mean (sd))	76.51 (19.71)	65.01 (26.47)	<0.001	0.493
PgR (mean (sd))	55.82 (32.22)	50.47 (31.61)	0.055	0.168
Ki67 (mean (sd))	17.69 (12.60)	22.24 (13.98)	<0.001	0.342
T1abc				
0	5 (1.9)	4 (1.5)		
1	58 (22.1)	46 (17.6)		
2	199 (76.0)	212 (80.9)		
Luminal (%)	138 (52.7)	172 (65.6)	0.003	0.266
Relapse (%)	13 (5.0)	18 (6.9)	0.459	0.081
LocalR (%)	6 (2.3)	6 (2.3)	1.000	0
DistantR (%)	8 (3.1)	11 (4.2)	0.640	0.061
Death (%)	2 (0.8)	4 (1.5)	0.681	0.072
obsT (mean (sd))	2184.97 (942.09)	2628.29 (1041.92)	<0.001	0.446

Differences existing in the original sample (see Table 1) between groups in age, tumor grade and dimension, hor- mone expression, and proliferation index are less apparent, but still statistically significant. This persisting lack of parity between treatment groups, even after matching, limits the ability to estimate causal treatment effects from the propensity score matched sample.

Table 4: Control and Treatment Group Means in Synthetic Minority Over-sampling (SMOTE) Sample. Matched sample characteristics of interest, grouped by treatment arm

	Endocrine Therapy	Chemothrapy + Endocrine Therapy	P-value	Std. Difference
n	1000	1000		
Age (mean (sd))	52.55(10.14)	52.93(9.22)	0.377	0.04
G (mean (sd))	2.28(0.57)	2.27(0.57)	0.459	0.033
ER (mean (sd))	72.79(23.25)	70.55(23.46)	0.032	0.096
PgR (mean (sd))	54.40(33.02)	52.41(31.18)	0.167	0.062
Ki67 (mean (sd))	20.71(14.42)	21.30 (13.59)	0.347	0.042
T1abc			0.165	0.009
0	9 (0.9)	18 (1.8)		
1	175 (17.5)	161 (16.1)		
2	816 (81.6)	821 (82.1)		
Luminal (%)	618(61.8)	626(62.6)	0.747	0.016
Relapse (%)	12(1.2)	76(7.6)	0.130	0.071
LocalR (%)	45(4.5)	34(3.4)	0.011	0.0119
DistantR (%)	96(9.6)	33(3.3)	0.204	0.062
Death (%)	12(1.2)	14(1.4)	0.844	0.018
obsT (mean (sd))	2252(983)	2643(1022)	<0.001	0.39

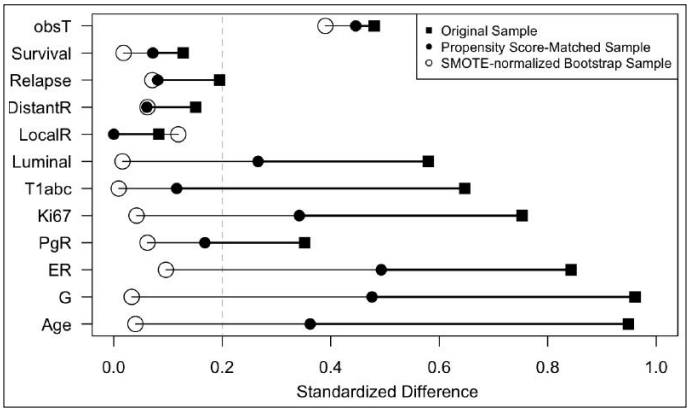

Figure 3: Standardized Differences Between Treatment and Control

As illustrated by Figure 3, the propensity score matching procedure has reduced the standardized differences of all the observed variables, bringing the difference between treatment arms closer to a hypothetical random sample that would exist in the event of a clinical trial. While several variables such as observation time, tumor grade, and dimensions have dropped to acceptable standardized differences as a result of the matching, most other variables remain well above acceptable limits for inter-group differences (standardized differences > 0.2 denoted by a grey, dashed line in Figure 3).

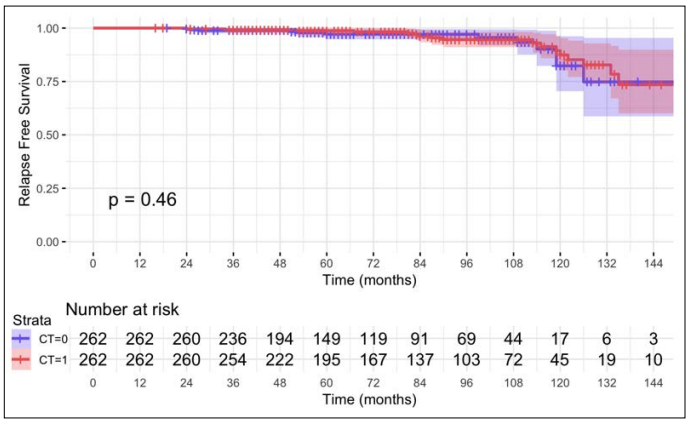

Figure 4: Kaplan Meier Curves for Matched Sample

Traditional Kaplan-Meier survival analysis of the propensity score-matched samples, as seen in Figure 5, shows no significant survival benefit of the treatment arm (Log-Rank Test; p-value=0.46). Both direct comparison of proportional relapse in the propensity-matched sample as well as Kaplan-Meier survival analysis suggest that, following surgical removal of a tumor, the addition of chemotherapy to a treatment regimen that includes hormonal therapy confers no (statistically significant) additional survival benefit for individuals with HR+, HER2-, T1, N0, M0 breast cancer under current treatment selection methodology. The two treatment groups included heavily imbalanced background characteristics, indicating that subjects were systematically sorted into the two groups, a non-random process that was difficult to control for in this survival analysis, even after PSM.

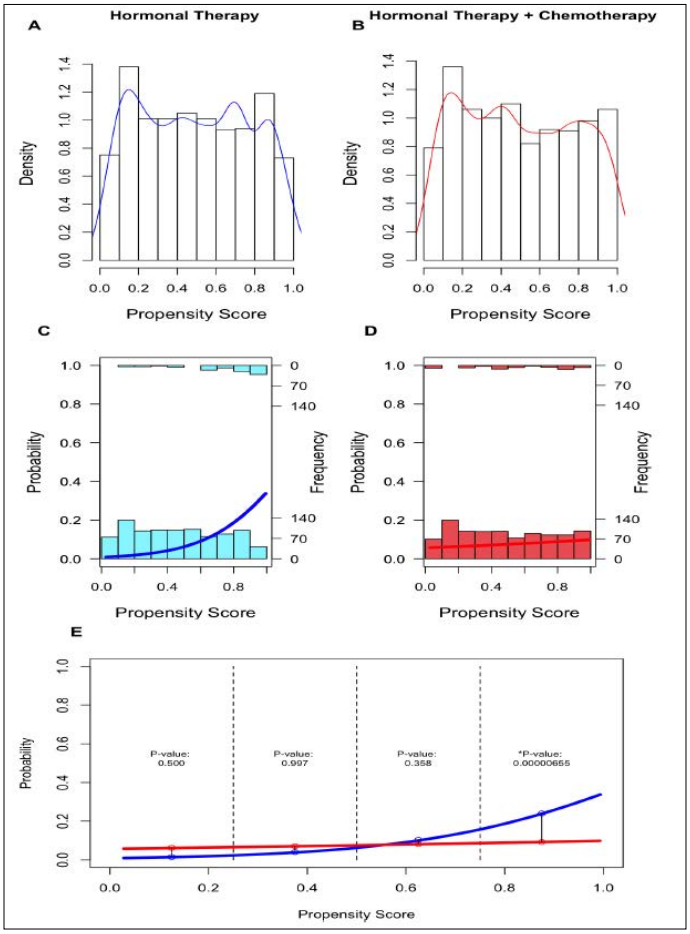

Figure 5: Bootstrapped Simulation (n=1000) using SMOTE normalization

Equation 2: The SMOTE normalization function:

$$P(X_i, d) = \frac{E[X_i]}{n} \int_{d}^{d+1} x\,dx \qquad (2)$$

Selection probability normalization with SMOTE was conducted by assigning a selection probability to each data point in the matched sample before bootstrapping. This probability was calculated using the inverse probability of a value for a given locale, here defined as a propensity score decile (d) and n is the sample size. E (Xi) is the average number of events across all datapoints, d and d+1 are the lower and upper boundaries of an arbitrary interval of support of a distribution. For a decile with below-average proportion of events, the normalized sampling probability will be inflated. For a decile with above-average proportion of events, the opposite is true.

In this study, the intervals are deciles of an uncharacterized distribution with support: S= [0,1], but the interval can be arbitrarily small as long as the number of intervals remains less than n. Additionally, the equation for determining inverse-frequency sample probabilities for SMOTE can be generalized to continuous distributions for Baysian applications.

A simulated dataset using a selection probability normalized, bootstrapped sample (n=1000) of the matched dataset revealed differences in relapse proportion between individuals' treatment arms with propensity scores in the highest quartile (Q4). No difference in proportion of relapse for Q1, Q2, and Q3 of propensity scores (Figure 5 E). Differences in proportions, for corresponding quartiles of the treatment and control groups, was tested using a 2-Sample t-test for proportions, with a cumulative alpha=0.05 and using a Bernoulli correction to account for multiple testing.

Discussion

The similarity of distributions and densities between Figure 5 A and B are evidence that the theoretical balancing of SMOTE has, in practice, worked in effectively eliminating all discernible differences between study arms. Bootstrapping and SMOTE sample distribution normalization do not change the within-group proportions, because the simulated groups in Figure 5 E are sampling from the same underlying distributions as the matched samples data in Figure 2.

The SMOTE-normalized sample did show a significant difference in proportion of relapse for the highest propensity quartile (p-value=6.55x10-6). This means, that individuals with propensity scores above 0.75, who were allocated to the endocrine monotherapy treatment, experienced a significantly higher rate of relapse. Given the near-equivalence of the two SMOTE sample treatment arms, the effect observed in these individuals should be attributed to the under- treatment. Conversely, since the relapse rates for individuals with PS<0.75 with chemoendocrine therapy is not statistically different from endocrine monotherapy. Given that the endocrine monotherapy causes fewer side effects, Figure 5 E indicates that there is widespread over-treatment of these individuals.

Conclusions

In order to create a prognostic index for HR+, HER2-, T1, N0, M0 breast cancer, in the absence of gene profiling technology, this paper statistically tests an existing, robust, empirical risk-quantifying model: physician intuition (modeled by a logit regression). Through sequential matching and synthetic resampling, PSM and SMOTE help differentiate cause and effect, by simulating a randomized clinical trial, in order quantify treatment effect against the aforementioned empirical regression.

Where advanced techniques of propensity score matching and survival analysis were unable to differential statistically significant survival benefit between SoC treatments of HR+, HER2-, T1, N0, M0 breast cancer, bootstrapped and SMOTE-normalized re-sampling simulations indicate strong evidence that, all else equal, there is statistically significant decrease in proportion of observed relapses, as a result of chemotherapy treatment, for individuals with PS > 0.75, treated with hormonal therapy in a large, in a randomized clinical trial of HR+, HER2-, T1, N0, M0 breast cancer. However, these findings must be understood to be provisional, as the study was observational in nature and the non-random nature of treatment allocation caused significant class imbalance.

The initial null findings of PSM and Kaplan-Meier analysis can be understood through examination of violated critical assumptions prerequisite to both statistical methods: causal effect estimation capabilities of propensity score matching and Kaplan-Meier survival analysis are only possible under the precondition that comparator groups can be assumed to be reasonably equal in distribution of, at a minimum, all known covariates, which was not the case for this dataset. As delineated in the statistical analysis plan, the initial findings were understood to be provisional, as the dataset was observational in nature, and the non-random nature of treatment allocation caused significant class imbalance.

The logistic regression used to generate propensity scores is an estimate of cumulative clinical decisions, risk of relapse quantification, made by physicians of individuals in the cohort (Table 2). This regression shows potential for use as prognostic index for risk of relapse in HR+, HER2-, T1, N0, M0 breast cancer. Through PSM and SMOTE, the statistical analysis detailed above was able to mathematically characterize this intuition, and statistically validate the efficacy of this intuition by simulating a scenario that holds all else equal in order to properly quantify true treatment effect (as is done in randomized clinical trials).

Statistical analysis in this paper strongly suggests a cutoff at PS=0.75, above which standard of care treatments differ significantly in risk to the patient. Above this point, chemoendocrine therapy is significantly more beneficial for the patient over endocrine monotherapy. Fortunately, post-hoc analysis indicates that, of the 1235 patients observed, only 13 were inadvertently put at higher risk of relapse.

However, the converse of this was much more widespread: T-test in Figure 5 E suggest that there is no expected benefit from the addition of chemotherapy to endocrine therapy for HR+, HER2-, T1, N0, M0 individuals with PS < 0.75 and thus should not be recommended by physicians below this cutoff. Retrospective analysis of the dataset includes 164 instances of over-treatment, as defined by the aforementioned criteria, where individuals were given additional chemotherapy unnecessarily.

Insight into both over-treatment and under-treatment in HR+, HER2-, T1, N0, M0 breast cancer, as well as definitive statistical findings outlined in this paper, indicate the potential for the logit regression equation, seen in Table 2, to be used as a clinical decision-making tool for defining optimal treatment regimens by reliably quantifying risk of relapse in the absence of gene profiling technology.

Acknowledgements

The anonymized clinical dataset was kindly supplied by the Department of Medical Oncology, University of Naples "Federico II".

Conflict of Interest

All authors declare no conflicts of interest.

References

1. Murphy C, Dickler M (2016) Endocrine resistance in hormone-responsive breast cancer: Mechanisms and therapeutic strategies. Endocr Relat Cancer 8: R337-R352.
2. Gamucci T, Vaccaro A, Ciancola F, Pizzuti L, Sperduti I, et al. (2013) Recurrence risk in small, node-negative, early breast cancer: A multicenter retrospective analysis. J Cancer Res Clin Oncol 5: 853-860.
3. O'Sullivan C, Loprinzi C, Haddad T (2018) Updates in the Evaluation and Management of Breast Cancer. Mayo Clinic Proceed- ings 6: 794-807.
4. Gradishar W, Salerno K (2016) NCCN guidelines update: Breast cancer. J Natl Compr Cancer Netw 4: 641-644.
5. Esserman L, Anton-Culver H, Borowsky (2017) WISDOM study and Athena Investigators. The WISDOM Study: breaking the deadlock in the breast cancer screening debate.

npj Breast Cancer 1: 34.

6. Curigliano G, Burstein H, P-Winer E, Gnant M, Dubsky P, et al. (2017) De-escalating and escalating treatments for early- stage breast cancer: The St. Gallen International Expert Consensus Conference on the Primary Therapy of Early Breast Cancer 2017. Ann Oncol 8: 1700-1712.
7. Mantel N (1966) Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemother 1: 163-170.
8. Harringtonm D, Flemingm T (1982) A Class of Rank Test Procedures for Censored Survival Data. Biometrika 69: 553-566.
9. Schoenfeld D (1981) The asymptotic properties of nonparametric tests for comparing survival distributions. Biometrika 68: 316-319.
10. Peto R, Peto J (1972) Asymptotically Efficient Rank Invariant Test Procedures. Journal of the Royal Statistical Society, Series A (General) 2: 185-207.
11. Kourou K, Exarchos T, Exarchos K, Karamouzis M, Fotiadis D (2015) Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J 1: 8-17.
12. Manor O, Segal E (2013) Predicting Disease Risk Using Bootstrap Ranking and Classification Algorithms. PLoS Comput Biol 8: e1003200.
13. Chawla N, Bowyer K, Hall L, Kegelmeyer W (2002) SMOTE: Synthetic minority over-sampling technique. J Artif Intell Res 1: 321-357.