



## Primary Site Surgery and Survival Impact in Metastatic HER2-Amplified Breast Cancer: Responsible use of Cohort Data

Robert W. Krell, MD<sup>1</sup> , and Philip M. Spanheimer, MD<sup>2</sup>

<sup>1</sup>Brooke Army Medical Center, Fort Sam Houston, TX; <sup>2</sup>University of North Carolina at Chapel Hill, Chapel Hill, NC

We read with interest the article by Mudgway et al. entitled “The Impact of Primary Tumor Surgery on Survival in HER2 Positive Stage IV Breast Cancer Patients in the Current Era of Targeted Therapy”.<sup>1</sup> With significant advances in systemic therapy for human epidermal growth factor receptor 2 (HER2)-amplified breast cancer, the therapeutic advantage of primary site resection in metastatic disease remains an important and open question.

Cohort data are the main sources for studying the role of surgery in metastatic disease for a variety of cancers. In the setting of metastatic colorectal cancer, for example, there are no randomized trial data comparing surgery with systemic therapy; however, based on the preponderance of evidence, surgery remains the preferred treatment in patients with resectable disease.<sup>2</sup> Nonetheless, we have more information regarding surgical selection and outcomes in patients with metastatic breast cancer, including several randomized controlled trials (RCT) and a Cochrane review.<sup>3–7</sup> The randomized trials have not provided a clear answer to this question; two have shown no difference in outcomes,<sup>3,6</sup> another suggested a survival advantage with primary site surgery,<sup>5</sup> and a fourth was closed early due to poor accrual.<sup>4</sup> Perhaps not surprisingly, the Cochrane review could not demonstrate a clear survival advantage with primary site surgery either.<sup>7</sup> It may be impractical to run an RCT powered to answer this question in the subset of patients with HER2-amplified breast cancer. This is

a data void that high-quality cohort studies can fill in. In their manuscript, Mudgway et al. conclude primary site surgery is associated with increased overall survival in patients with metastatic HER2-amplified breast cancer. Their results mirror other authors using the National Cancer Data Base and using similar techniques,<sup>8</sup> but contrast with other studies using different data sets and techniques,<sup>9</sup> as well as published randomized trials. As with other tough clinical questions, we are left wondering which technique is best to address the question.

Cohort data provide large numbers to overcome study power limitations. For the researcher and the reviewer, the main drawback to cohort studies is that treatments, especially in the setting of metastatic disease, are subject to considerable selection bias.<sup>10–12</sup> To address this are an armamentarium of statistical analytic techniques that account for overt and unmeasured bias, including hierarchical modeling, cohort restriction and sensitivity analyses, propensity scores, and econometric techniques. In the present paper, the authors use multivariable regression with propensity score matching as a type of sensitivity analysis. While both are appropriate approaches to the clinical question, the particulars of their application in this study, and therefore the validity of their outputs, is unknown.

Propensity score techniques are a relatively recent addition to the surgical literature, but are widely used in other settings.<sup>11–13</sup> In essence, a subject’s probability of experiencing an exposure of interest is modeled using available covariates in the data; the resulting probability is that subject’s propensity score, which can then be used as an additional adjustment covariate or used to match patients with similar characteristics in a quasi-randomized experiment. Propensity score models and techniques are valuable tools for assessing cohort data, and are available in most commercially available statistical packages; however, their ease of employment undercuts their responsible

---

© This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2020

First Received: 29 June 2020  
Accepted: 4 July 2020;  
Published Online: 22 July 2020

R. W. Krell, MD  
e-mail: robwkrell@gmail.com

use. Most propensity scores are derived using probit or logistic regression techniques, and, like any other regression, the final product (the propensity score) depends on the quality of the model used to generate it: the covariates used, their relationship with the exposure, and the model's underlying assumptions and fit, etc. Several studies have shown that when applied indiscriminately, the outputs from propensity models generally yield similar results (in terms of treatment effect estimates) as 'conventional' multivariable regression.<sup>14</sup> Most importantly, the models used to derive propensity scores are subject to the same biases as conventional regression models. In other words, the same 'garbage-in-garbage-out' principles for multivariable analyses also apply to propensity score analyses. For example, it is obvious from the published trial data that response to systemic therapy and metastatic burden have strong influences on patient selection for surgery.<sup>3,5,6</sup> In study by Mudgway et al., the lack of metastatic burden and disease response information in the data are important, but even more important is a noticeable lack of techniques used to account for those and other missing data. Their methods are too opaque for the reader to assess selection bias mitigation in their study.

With the increase in propensity score modeling reported in the medical literature, there is a clear need for researchers to be transparent with their methods to facilitate peer review. In fact, there are several publicly available best practice guidelines, some specific to cancer research, that address this issue.<sup>15–17</sup> All agree that authors should state their rationale for selecting their statistical methods and model variables, how they address missing data in their cohort, diagnostics used for assessing the regression model, and how adjustment/propensity scores change treatment effects from the baseline cohort. Unfortunately for the reader, the details of the present study's analyses are not shown, their models assumptions and performance left unstated, and the treatment effects unclear since they mention both inverse-probability weighting and adjustment after propensity matching. The reader is left with the conclusion that patients selected for primary site surgery do better than those treated with systemic therapy alone, which is contradicted by RCTs. What remains unclear is whether their methodology was strong enough to underscore their conclusion that in the subset of HER2-enriched patients, RCT data should be disregarded in favor of their cohort data.

The authors use the word 'impact' in the title, implying a causal link between surgery and survival in this setting implies. This wording is problematic because the majority of physicians and surgeons treating patients with breast cancer are not trained extensively in statistical methods to independently develop a nuanced understanding of the limitations of these data. It is easy to imagine a situation

where someone would conclude, based on the study's title and presentation, that primary site surgery should be offered to patients with metastatic HER2-amplified breast cancer, and counsel the patient that they are expected to live longer with that approach. This is not supported by the data and is a potentially dangerous strategy. To be clear, we do not advocate for withholding primary site surgery in metastatic disease, rather we advocate that discussions on the matter should not cite survival benefits that are unproven in RCTs and are unsubstantiated by opaque analyses of cohort data.

Surgery in the setting of metastatic disease is an exercise in patient selection, and we as a surgical community have generally been excellent in selecting the right patients. However, to derive the most accurate estimate of surgery's value (and not the surgeon's selection) we need carefully conducted and honestly reported cohort studies.

**FUNDING** None

**DISCLOSURES** The views expressed are those of the author and do not reflect the official policy or position of the US Air Force, Department of Defense, or the US Government.

## REFERENCES

1. Mudgway R, Chavez de Paz Villanueva C, Lin AC, Senthil M, Garberoglio CA, Lum SS. The impact of primary tumor surgery on survival in HER2 positive stage IV breast cancer patients in the current era of targeted therapy. *Ann Surg Oncol*. 2020;27:2711–2720. <https://doi.org/10.1245/s10434-020-08310-2>.
2. National Comprehensive Cancer Network. NCCN guidelines: colon cancer, version 4.2020. 2020. Available at: <https://www.nccn.org>. Accessed 20 Jun 2020.
3. Badwe R, Hawaldar R, Nair N, et al. Locoregional treatment versus no treatment of the primary tumour in metastatic breast cancer: an open-label randomised controlled trial. *Lancet Oncol*. 2015;16:1380–8.
4. Fitzal F, Bjelic-Radisic V, Knauer M, et al. Impact of breast surgery in primary metastasized breast cancer: outcomes of the prospective randomized phase III ABCSG-28 POSYITIVE Trial. *Ann Surg*. 2019;269:1163–9.
5. Soran A, Ozmen V, Ozbas S, et al. Randomized trial comparing resection of primary tumor with no surgery in stage IV breast cancer at presentation: protocol MF07-01. *Ann Surg Oncol*. 2018;25:3141–9.
6. Khan SA, Zhao F, Solin LJ, et al. A randomized phase III trial of systemic therapy plus early local therapy versus systemic therapy alone in women with de novo stage IV breast cancer: a trial of the ECOG-ACRIN research group (E2108). *J Clin Oncol*. 2020;38(18 Suppl):LBA2.
7. Tosello G, Torloni MR, Mota BS, Neeman T, Riera R. Breast surgery for metastatic breast cancer. *Cochrane Database Syst Rev*. 2018;3:CD011276.
8. Lane WO, Thomas SM, Blitzblau RC, et al. Surgical resection of the primary tumor in women with de novo stage IV breast cancer: contemporary practice patterns and survival analysis. *Ann Surg*. 2019;269:537–44.

9. Dominici L, Najita J, Hughes M, et al. Surgery of the primary tumor does not improve survival in stage IV breast cancer. *Breast Cancer Res Treat.* 2011;129:459–65.
10. Giordano SH, Kuo Y-F, Duan Z, Hortobagyi GN, Freeman J, Goodwin JS. Limits of observational data in determining outcomes from cancer therapy. *Cancer.* 2008;112:2456–66.
11. Bosco JLF, Silliman RA, Thwin SS, et al. A most stubborn bias: no adjustment method fully resolves confounding by indication in observational studies. *J Clin Epidemiol.* 2010;63:64–74.
12. Joffe MM, Rosenbaum PR. Invited commentary: propensity scores. *Am J Epidemiol.* 1999;150:327–33.
13. Grose E, Wilson S, Barkun J, et al. Use of propensity score methodology in contemporary high-impact surgical literature. *J Am Coll Surg.* 2020;230:101–12.
14. Stürmer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. *J Clin Epidemiol.* 2006;59:437.e1–e24.
15. Institute of Social and Preventive Medicine. 2009. STROBE Statement: strengthening the reporting of observational studies in epidemiology. Available at: <https://www.strobe-statement.org>. Accessed 22 Mar 2020.
16. McMurry TL, Hu Y, Blackstone EH, Kozower BD. Propensity scores: methods, considerations, and applications in the Journal of Thoracic and Cardiovascular Surgery. *J Thorac Cardiovasc Surg.* 2015;150:14–19.
17. Yao XI, Wang X, Speicher PJ, et al. Reporting and guidelines in propensity score analysis: a systematic review of cancer and cancer surgical studies. *J Natl Cancer Inst.* 2017;109:djw323.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.