

## SHORT COMMUNICATION

# Machine learning with PROs in breast cancer surgery; caution: Collecting PROs at baseline is crucial

Laurentine S. E. van Egdom MD<sup>1</sup> | Andrea Pusic MD, MHS, FACS, FRCSC<sup>2</sup> |  
Cornelis Verhoef MD, PhD<sup>1</sup> | Jan A. Hazelzet MD, PhD<sup>3</sup> |  
Linetta B. Koppert MD, PhD, MSc<sup>1</sup>

<sup>1</sup>Department of Surgical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

<sup>2</sup>Department of Plastic and Reconstructive Surgery, Patient-Reported Outcomes, Value & Experience (PROVE) Center, Brigham and Women's Hospital, Boston, MA, USA

<sup>3</sup>Department of Public Health, Erasmus MC, University Medical Center, Rotterdam, The Netherlands

## Correspondence

Laurentine S. E. van Egdom, Department Of Surgical Oncology, Erasmus MC Cancer Institute, Dr Molenwaterplein 40, 3000 CA Rotterdam, The Netherlands.  
Email: l.vanegdom@erasmusmc.nl

## Funding information

Vereniging Trustfonds Erasmus Universiteit Rotterdam

## Abstract

As high breast cancer survival rates are achieved nowadays, irrespective of type of surgery performed, prediction of long-term physical, sexual, and psychosocial outcomes is very important in treatment decision-making. Patient-reported outcomes (PROs) can help facilitate this shared decision-making. Given the significance of more personalized medicine and the growing trend on the application of machine learning techniques, we are striving to develop an algorithm using machine learning techniques to predict PROs in breast cancer patients treated with breast surgery. This short communication describes the bottlenecks in our attempt to predict PROs.

## KEYWORDS

breast cancer surgery, machine learning, patient-reported outcomes

Improvement in early detection and treatment of breast cancer has resulted in increased long-term breast cancer survivors.<sup>1</sup> The cornerstone of breast cancer management still is surgery. In breast cancer surgery, equal survival rates are achieved, irrespective of type of surgery performed.<sup>2-4</sup> However, breast cancer surgery can adversely affect women's psychological health and health-related quality of life (HRQoL). Prediction of long-term physical, sexual, and psychosocial outcomes is therefore very important in treatment decision-making.

Patient-reported outcomes (PROs) come directly from the patient without interpretation by a health care provider and reflect aspects of health, quality of life, and related constructs.<sup>5</sup> The routine collection of PROs has been implemented in many health institutions,<sup>6-10</sup> and it is clear that PROs have an important role in today's clinical practice. Collaboration of the International Consortium for Health Outcomes Measurement (ICHOM) with several other health

care institutions worldwide has resulted in the development of a Standard Set for breast cancer outcomes.<sup>10</sup> Within this outcome set, patient-reported outcome measures (PROMs) are pivotal and accounting for 75% of the outcomes evaluated.<sup>10</sup>

PROs can help facilitate in shared decision-making through informing treatment decisions and setting expectations. The ability for patients to understand what other patients with breast cancer experienced after surgery is thereby vital.

Predictive modeling is not new to medicine. In clinical medicine, a multivariable prediction model combines information from multiple predictors to predict the probability of or risk for a specific disease or outcome.<sup>11</sup> Predictive modeling has the purpose of informing patients and guiding clinicians in decision-making on treatment decisions. The majority contains prediction of patient outcomes focused on cancer survival and risk of cancer recurrence/local control,<sup>12-14</sup>

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *The Breast Journal* published by Wiley Periodicals, Inc.

but little has been done to predict PROs into the future. Moreover, to our knowledge, there are no tools available focusing on predicting HRQoL outcomes after breast surgery into the future. Given the significance of more personalized medicine and the growing trend on the application of machine learning techniques, our breast cancer team is striving to develop an algorithm using machine learning techniques to predict PROs in breast cancer patients treated with breast surgery.

We aimed to develop and validate a simple prediction model for improvement of HRQoL after breast cancer surgery using data from three PRO questionnaires as proposed in the ICHOM Standard Set for Breast Cancer, namely the EORTC QLQ-C30 and EORTC QLQ-B23, and the BREAST-Q (postoperative modules). To this end, a retrospective cohort collected and described previously<sup>6</sup> was used. This cohort contained 764 female patients with breast cancer (pTis-3N0-3M0) who underwent breast cancer surgery between January 2005 and September 2016 at Erasmus MC Academic Breast Cancer Center, Rotterdam, the Netherlands. Data on patient characteristics, age, date and type of surgery, tumor morphology, TNM staging (7th edition<sup>15</sup>), hormonal status, HER2 status, BRCA 1/2 gene mutation status, local recurrence, second primary breast cancer, details regarding chemotherapy and/or immunotherapy and endocrine therapy, radiotherapy, and follow-up were available. Machine learning (ie, general linear model regression (GLM), support vector machines (SVM), single-layer artificial neural networks (ANN), and deep learning (DL))<sup>16</sup> was used to jointly study presurgical prognostic variables relating to age, medical status, tumor characteristics, and possible (neo)adjuvant treatment indications/treatment characteristics. Unfortunately, a lack of relationship was found between outcome variables and their predictors, meaning that the accuracy reflected just the population prevalence of the outcomes. Machine learning models have an immense number of parameters that must be either learned using data or set manually by the researcher.<sup>17</sup> By combining variables in a reduced number of dimensions, we tried to help the analysis, but this did not yield substantial changes and required days of computational time.

During the process, some crucial obstacles were identified, which stagnated the development of a machine learning model in this dataset. This included the cross-sectional design, the lack of baseline PROs, and the relative small sample size. Given the increase in the use of machine learning techniques in medical research and the, worldwide, desire to predict and influence PROs after breast surgery, we believe it is important to draw attention to our findings.

Machine learning describes the use of computer algorithms that learn nonlinear associations retrospectively from the data to estimate risk of a specific outcome. Even though machine learning is increasingly used in medical research,<sup>18-20</sup> success is not always guaranteed. As with any method, a good understanding of the problem and an appreciation of the limitations of the dataset is important. Also crucial is an understanding of the assumptions and limitations of the algorithms being applied. If a machine learning experiment

is properly designed, with correct implementation and validated results, there usually is a good chance of success.

Although we used patient and treatment characteristics, and outcomes of interest to both patients and clinician (ie, validated PROMs as proposed in the ICHOM Standard Set for Breast Cancer), there were some important limitations in using the existing dataset.<sup>6</sup> With 764 breast cancer patients, the study was relatively large, although for machine learning techniques probably not large enough. The size of the dataset is one of the most common limitations noted in studies reporting machine learning techniques.<sup>14</sup> The dataset needs to be sufficiently large, which allows sufficient partitioning into training and testing sets, leading to reasonable validation of the estimators<sup>14</sup> in order to enhance the generalizability of the predictive model.

The most important limitation however is the cross-sectional design of the dataset, meaning the absence of baseline PROs. Traditional methods for evaluating PROMs look at the change over time, using the baseline compared with the end point. Enabling comparison with preoperative PROs is expected to reflect the influence of different treatments on HRQoL outcomes better than a single score obtained following treatment. One explanation probably is the fact that not every individual patient will score their breasts to the highest possible level at baseline. Although preoperative PROs were not available, all known other potential predictors were assessed, except for socioeconomic status (which cannot be easily obtained in the Netherlands for privacy reasons). The next step toward further validation of this approach to prediction would be to work with a more complete dataset, including baseline PROs and lifestyle measures. The research team has secured a prospective dataset over a longer time frame, but this dataset currently consists of a small number of patients. Since PROM collection is considered standard of care at our institute nowadays,<sup>9</sup> in combination with a regional and international collaboration, this cohort will be progressively enlarged over time. There are plans in place to develop and test the performance of the machine learning techniques in this dataset in the near future. However, the above-described study was a valuable first step toward modeling PROM data for use in breast cancer surgery. Once developed, the model could have potential for use outside breast surgery because similar sets are used in other diseases. But, as also suggested by Beam et al,<sup>17</sup> the challenges and obstacles to reproducibility of machine learning techniques must be carefully considered to ensure the validation, safety, and effectivity of these new class of prediction tools.

In conclusion, using machine learning methods, we endeavored to develop a clinical prediction model for PROs after breast surgery. Clinicians could use information on the level of patient HRQoL outcome improvement, when counseling patients about the (prognostic) outcomes of breast cancer surgery, allowing patients to be more involved in their treatment decision. To actually realize an effective clinical prediction model, information regarding patients' starting position is crucial. This emphasizes the urgent need of collecting PROMs at baseline, leading to the opportunity of predictive modeling on PROMs in breast cancer surgery in the future.

## REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68(6):394-424.
2. Lagendijk M, van Maaren MC, Saadatmand S, et al. Breast conserving therapy and mastectomy revisited: breast cancer-specific survival and the influence of prognostic factors in 129,692 patients. *Int J Cancer.* 2018;142(1):165-175.
3. Fisher B, Anderson S, Bryant J, et al. Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. *N Engl J Med.* 2002;347(16):1233-1241.
4. Litiere S, Werutsky G, Fentiman IS, et al. Breast conserving therapy versus mastectomy for stage I-II breast cancer: 20 year follow-up of the EORTC 10801 phase 3 randomised trial. *Lancet Oncol.* 2012;13(4):412-419.
5. Browne JP, Cano SJ, Smith S. Using patient-reported outcome measures to improve health care: time for a new approach. *Med Care.* 2017;55(10):901-904.
6. Lagendijk M, van Egdom LSE, van Veen FEE, et al. Patient-reported outcome measures may add value in breast cancer surgery. *Ann Surg Oncol.* 2018;25(12):3563-3571.
7. Howell D, Molloy S, Wilkinson K, et al. Patient-reported outcomes in routine cancer clinical practice: a scoping review of use, impact on health outcomes, and implementation factors. *Ann Oncol.* 2015;26(9):1846-1858.
8. Chen J, Ou L, Hollis SJ. A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Serv Res.* 2013;13:211.
9. van Egdom LSE, Lagendijk M, van der Kemp MH, et al. Implementation of value based breast cancer care. *Eur J Surg Oncol.* 2019;45(7):1163-1170.
10. Ong WL, Schouwenburg MG, van Bommel ACM, et al. A standard set of value-based patient-centered outcomes for breast cancer: the international consortium for health outcomes measurement (ICHOM) initiative. *JAMA Oncol.* 2017;3(5):677-685.
11. Moons KG, Royston P, Vergouwe Y, et al. Prognosis and prognostic research: what, why, and how? *BMJ.* 2009;338:b375.
12. Quinten C, Maringwa J, Gotay CC, et al. Patient self-reports of symptoms and clinician ratings as predictors of overall cancer survival. *J Natl Cancer Inst.* 2011;103(24):1851-1858.
13. Osoba D. Health-related quality of life and predicting survival in cancer: not a simple matter. *Support Care Cancer.* 2007;15(4):353-355.
14. Kourou K, Exarchos TP, Exarchos KP, et al. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J.* 2015;13:8-17.
15. Sobin LH, Gospodarowicz M, Wittekind C. *TNM classification of malignant tumors.* New York, NY: John Wiley & Sons, Inc.; 2009.
16. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. *BMC Med Res Methodol.* 2019;19(1):64.
17. Beam AL, Manrai AK, Ghassemi M. Challenges to the reproducibility of machine learning models in health care. *JAMA.* 2020;323(4):305.
18. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med.* 2018;24(9):1342-1350.
19. Huang C, Murugiah K, Mahajan S, et al. Enhancing the prediction of acute kidney injury risk after percutaneous coronary intervention using machine learning techniques: a retrospective cohort study. *PLoS Med.* 2018;15(11):e1002703.
20. Nanayakkara S, Fogarty S, Tremeeer M, et al. Characterising risk of in-hospital mortality following cardiac arrest using machine learning: a retrospective international registry study. *PLoS Med.* 2018;15(11):e1002709.

**How to cite this article:** van Egdom LSE, Pusic A, Verhoef C, Hazelzet JA, Koppert LB. Machine learning with PROs in breast cancer surgery; caution: Collecting PROs at baseline is crucial. *Breast J.* 2020;26:1213-1215. <https://doi.org/10.1111/tbj.13804>