

MINI REVIEW

Open Access



Are online prediction tools a valid alternative to genomic profiling in the context of systemic treatment of ER-positive breast cancer?

Umar Wazir^{*} , Kinan Mokbel, Amtul Carmichael and Kefah Mokbel

^{*} Correspondence:
umar.wazir@rcsed.ac.uk
The London Breast Institute,
Princess Grace Hospital, 45
Nottingham Place, London W1U
5NY, UK

Abstract

Background: Clinicians use clinical and pathological parameters, such as tumour size, grade and nodal status, to make decisions on adjuvant treatments for breast cancer. However, therapeutic decisions based on these features tend to vary due to their subjectivity. Computational and mathematical algorithms were developed using clinical outcome data from breast cancer registries, such as Adjuvant! Online and NHS PREDICT. More recently, assessments of molecular profiles have been applied in the development of better prognostic tools.

Methods: Based on the available literature on online registry-based tools and genomic assays, we evaluated whether these online tools could be valid and accurate alternatives to genomic and molecular profiling of the individual breast tumour in aiding therapeutic decisions, particularly in patients with early ER-positive breast cancer.

Results and conclusions: Early breast cancer is currently considered a systemic disease and a complex ecosystem with behaviour determined by the complex genetic and molecular signatures of the tumour cells, mammary stem cells, microenvironment and host immune system. We anticipate that molecular profiling will continue to evolve, expanding beyond the primary tumour to include the tumour microenvironment, cancer stem cells and host immune system. This should further refine therapeutic decisions and optimise clinical outcome.

This article was specially invited by the editors and represents work by leading researchers.

Keywords: Breast cancer, Molecular biotechnology, Genomic assay, Prognosis, Gene expression regulation

Introduction

Traditionally, clinicians use clinical and pathological parameters, such as tumour size, grade, nodal status, HER2, ER status and proliferation index, to make decisions regarding adjuvant treatments for breast cancer. However, therapeutic decisions based on these features tend to vary due to their subjectivity [1, 2].

The modern approach uses computational and mathematical algorithms that were developed using clinical outcome data from cancer registries. Adjuvant! Online and

NHS PREDICT are examples of such decision-supporting tools. They are freely available online, making them attractive in resource-constrained healthcare settings. They help clinicians to assess an individual's risk of developing recurrent disease and/or dying within 10 years, and have the potential to guide decisions regarding adjuvant and neoadjuvant therapy [3, 4].

The deepening understanding of breast cancer has been used to significantly improve these types of prognostic tool. A particularly influential discovery was the characterisation of breast cancer as a heterogeneous group of neoplastic processes arising from the ductal or lobular epithelium rather than a single disease with a variable ER and HER2 expression. It enabled the development of better prognostic tools based on assessing molecular profiles. Examples of such assays include Blueprint [5] Mammaprint [6], Oncotype DX [7], prediction analysis of microarray 50 (PAM50) [8, 9] and EndoPredict (EP) [10]. The EndoPredict Clinical (EpClin) assay is a composite of standard pathological parameters and molecular profiling scores which has been found to provide superior prognostic information [11]. These assays have changed the landscape of clinical oncology and allowed clinicians to make therapeutic decisions based on the molecular machinery of the tumour and data derived from randomised controlled trials.

These commercially available molecular scores have not only been found to be cost-effective, they have become less expensive over the past few years. Despite this, their cost may be an issue in resource-poor settings. An analysis by Reed et al. suggests that the initial outlay on genomic assays are offset by future gains in quality-adjusted patient years [12]. However, cost remains a significant consideration.

In this article, we shall discuss the literature on the online tools mentioned above with a view to evaluating whether they could be valid and accurate alternatives to genomic and molecular profiling of the individual breast tumour in aiding therapeutic decisions in the era of personalized precision medicine, particularly in patients with early ER-positive breast cancer.

Online prediction tools

The online tools referred to earlier primarily use clinicopathological variables and cancer registry data as the basis of risk prediction. The clinical pathological variables used include age, tumour size and grade, mode of detection, number of lymph nodes involved, ER status, HER2 status, Ki67 status and type of chemotherapy [13]. The strengths and weaknesses of these tools draw from the design and limitations of the registry data on which they are based.

Adjuvant! Online

The baseline risk estimation for Adjuvant! Online was derived from the SEER (surveillance, epidemiology and end results) database program, which is a collation of nine databases covering 14% of the US population [14]. The SEER database specifically excludes patients under the age of 35 and over the age of 59 [15] and has limited information on the socio-economic status of subject. There have been concerns regarding the quality of the data about cause of death [16, 17].

The database does not include information regarding the benefits of adjuvant trastuzumab, thereby reducing the utility of Adjuvant! Online in clinical decisions about HER2-positive disease treatment [16, 17]. This deficiency of Adjuvant! Online with

regards to HER2-positive disease has significant implications for the prediction of metastatic spread. In a recent *in vitro* study using murine models, the HER2 status of cells predicted the response to progesterone-induced signalling, with HER2-deficient cells being more likely to migrate and HER2-enriched cells tending towards increased proliferation [18]. This recent evidence underlines the importance of HER2 in predicting prognosis and highlights the significance of this inherent shortcoming in online cancer registry-based prognostic tools.

Adjuvant! Online tends to overestimate the number of patients at high risk. Cardoso et al. found that Adjuvant! Online incorrectly classified 23% of patients as high clinical risk when Oncotype DX classified them as low genomic risk. [19].

In a population-based validation study, Olivotto et al. suggest that in patients under 35 years of age and who test positive for lymphovascular invasion, Adjuvant! Online would overestimate survival. It was also found that Adjuvant! Online tends to overestimate the survival rates of younger women with ER-positive breast cancer [3] and that it overestimated the added value of chemotherapy for older patients [20–22]. The validity of the predictive score calculated by Adjuvant! Online was deemed weak in the clinician-based validation [23]. Predictions on loco-regional relapse and distant metastases may vary greatly, making it difficult to make clear recommendations for adjuvant treatment [24]. This is reflected in two studies that suggest that when patients are involved in a discussion to decide on adjuvant chemotherapy, they are less likely to choose chemotherapy if using Adjuvant! Online [25, 26].

The ethnic variation in the data on which these online tools are based seriously affects the generalisability of these online tools. The SEER database is representative of the usual US population in terms of age, sex and ethnic distribution. However, the ethnic mix of the US population is different from that of England and Wales.

NHS PREDICT

The NHS PREDICT online tool is based on a cancer registry database of 5694 patients in the UK [4]. Unfortunately, independent investigators have raised concerns regarding the quality of the cancer registry data [27, 28]. Joishy et al. identified the lack of education of medical professionals and imprecision and inconsistency in medical records as factors negatively impacting the reliability of the data, and stated that insufficient time, personnel and finances had been allocated to ensure high quality [29]. The NHS PREDICT tool does not provide any estimate of local relapse and does not consider mortality due to causes other than breast cancer in its survival estimate. Therefore, a total reliance on the NHS PREDICT online tool may deprive patients, particularly those with small, biologically aggressive cancers, of the benefit of chemotherapy [24].

In our unpublished audit of 120 patients who underwent genomic profiling using both the EP Clin score and NHS PREDICT calculation, the discordance rate was significantly high (43%). If we relied solely on NHS PREDICT, a significant proportion of patients with small, node-negative tumours would not have received chemotherapy, despite needing adjuvant therapy according to the EP Clin score.

Wong et al. found that NHS PREDICT substantially overestimates survival in very young patients with breast cancer and those receiving chemotherapy [30].

As with Adjuvant! Online, the ethnic mix of the outcome data used to develop NHS PREDICT may not be generalisable to more diverse metropolitan areas in the UK, such as London and Birmingham.

In summary, online prediction tools continue to be of value as free adjuncts to therapeutic decision-making. However, the use of these tools should be tempered with recognition of the inherent biases of the underlying databases and the well-documented limitations of these algorithms, such as overestimation of the benefits of chemotherapy in certain patient groups, underestimation of the benefits of chemotherapy in patients with small, biologically aggressive tumours, lack of generalisability to more diverse populations, and lack of standardisation in clinical utility.

Genomic assays

The use of genomic assays in human breast cancer has been endorsed by the National Institute for Clinical Excellence (NICE) and the American Society of Clinical Oncology (ASCO), among others [31]. Several histological and molecular markers are used to identify patients with breast cancer at the highest risk of recurrence, which also means tumours with the highest degree of sensitivity to chemotherapy [11].

The most established example of a genomic assay score is the recurrence score (RS), which is based on an Oncotype DX 21-gene assay panel. The RS ranges from 1 to 100, and is stratified into low risk (below 18), intermediate risk (18 to 30) and high risk (over 30). RS was the first such score to be included in NICE and ASCO guidelines [31, 32].

EndoPredict (EP) is a 15-point score based on an 8-gene panel which assigns patients to high- and low-risk groups (below and above 5, respectively). The reliability of the score is increased by incorporating clinical parameters (tumour size and nodal status) in a score that has been named EndoPredict Clinical (EpClin). EP and EpClin have been shown to provide more prognostic information than RS, in part due to the combination of genomic data with nodal status and tumour size. The absence of an intermediate risk group in EP makes decision-making more straightforward than with RS [33]. EpClin provides reliable information about the benefit of chemotherapy by combining molecular signatures with clinicopathological variables and data derived from randomised controlled trials [34].

Mammaprint is the oldest available test, consisting of a 70-gene assay that stratifies the patients into high- and low-risk groups [19]. A further 80-gene panel called Blueprint was developed for more accurate typing of breast cancer (5). It is meant to be used in conjunction with Mammaprint. Mammaprint is still waiting for external validation in the MINDACT trial, the results of which were presented at ASCO [35].

PAM50 is a 50-gene assay that identifies the breast cancer subtype, and generates a risk of recurrence score (ROR). The ROR is a 100-point scale that stratifies patients into low risk of recurrence, intermediate risk and high risk. PAM50 was validated in the ATAC and ABCSG-8 trials, where it was found to be superior to immunohistochemistry and RS in ER-positive node-negative breast cancer patients receiving endocrine therapy [8, 9].

The efficacy of genomic assays is a testament to the milestones achieved in the understanding of cancer biology and the recent recognition of the heterogeneity of breast cancer as a disease.

Internet-based mathematical and computational algorithms provide physicians and patients with useful information regarding prognosis and the benefits of systemic therapy. They are particularly valuable in resource-constrained international healthcare. However,

their inherent limitations, which are related to the conceptual design, methodology and data quality, make these decision aids insufficiently robust to be used as an alternative to molecular profiling of the primary tumour in the modern era of personalised cancer care and precision medicine.

Early breast cancer is currently considered a systemic disease and a complex ecosystem with behaviour determined by complex genetic and molecular signatures of the tumour cells, mammary stem cells, microenvironment and host immune system rather than an anatomical neoplastic process that progresses locally and then spreads to regional lymph nodes and other organs during tumour progression. Therefore, we anticipate that molecular profiling will continue to evolve and expand to include the tumour microenvironment, cancer stem cells and host immune system, in addition to the primary tumour, to further refine therapeutic decisions and optimise clinical outcomes [36].

However, it would be prudent to closely follow developments in this field, as no single multi-gene assay is emerging as standard and no one technology is uniformly accepted. Continued studies for validation and reproducibility of genomic tests are needed to better understand their limitations and to further increase their utility in making treatment decisions in the early stages of breast cancer.

Acknowledgements

Not applicable.

Funding

This study was supported by funding from the Breast Cancer Hope Foundation (UK).

Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analysed during the study.

Authors' contributions

UW, KM, AC, and KM, AC and KM performed the literature review and proofread the article. UW and KM drafted the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 23 February 2017 Accepted: 21 August 2017

Published online: 04 September 2017

References

1. Haybittle JL, Blamey RW, Elston CW, et al. A prognostic index in primary breast cancer. *Br J Cancer*. 1982;45(3):361–6. PubMed PMID: 7073932. Pubmed Central PMCID: 2010939. Epub 1982/03/01. eng
2. Aaltomaa S, Lipponen P, Eskelinen M, et al. Mitotic indexes as prognostic predictors in female breast cancer. *J Cancer Res Clin Oncol*. 1992;118(1):75–81. PubMed PMID: 1729263
3. Olivetto IA, Bajdik CD, Ravdin PM, et al. Population-based validation of the prognostic model ADJUVANT! for early breast cancer. *J Clin Oncol*. 2005;23(12):2716–25. PubMed PMID: 15837986. Epub 2005/04/20. eng
4. Wishart GC, Azzato EM, Greenberg DC, et al. PREDICT: a new UK prognostic model that predicts survival following surgery for invasive breast cancer. *Breast Cancer Res*. 2010;12(1):R1. PubMed PMID: 20053270. Pubmed Central PMCID: 2880419. Epub 2010/01/08. eng
5. Whitworth P, Stork-Sloots L, de Snoo FA, et al. Chemosensitivity predicted by Blueprint 80-gene functional subtype and MammaPrint in the Prospective Neoadjuvant Breast Registry Symphony Trial (NBRST). *Ann Surg Oncol*. 2014;21(10):3261–7. PubMed PMID: 25099655. Pubmed Central PMCID: PMC4161926

6. Bueno-de-Mesquita JM, Linn SC, Keijzer R, et al. Validation of 70-gene prognosis signature in node-negative breast cancer. *Breast Cancer Res Treat.* 2009;117(3):483–95. PubMed PMID: 18819002. Epub 2008/09/27. eng
7. 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.* 2006;24(23):3726–34. PubMed PMID: 16720680. Epub 2006/05/25. eng
8. Dowsett M, Sestak I, Lopez-Knowles E, et al. Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. *J Clin Oncol.* 2013;31(22):2783–90. PubMed PMID: 23816962. Epub 2013/07/03. eng
9. Gnant M, Filipits M, Greil R, et al. Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. *Ann Oncol.* 2014;25(2):339–45. PubMed PMID: 24347518. Epub 2013/12/19. eng
10. Dubsky P, Filipits M, Jakesz R, et al. EndoPredict improves the prognostic classification derived from common clinical guidelines in ER-positive, HER2-negative early breast cancer. *Annals Oncol.* 2013;24(3):640–7. PubMed PMID: 23035151. Pubmed Central PMCID: 3574544. Epub 2012/10/05. eng
11. Wazir U, Mokbel K. Emerging gene-based prognostic tools in early breast cancer: first steps to personalised medicine. *World J Clin Oncol.* 2014;5(5):795–9. PubMed PMID: 25493218. Pubmed Central PMCID: 4259942. Epub 2014/12/11. eng
12. Reed SD, Dinan MA, Schulman KA, Lyman GH. Cost-effectiveness of the 21-gene recurrence score assay in the context of multifactorial decision making to guide chemotherapy for early-stage breast cancer. *Genet Med.* 2013;15(3):203–11. PubMed PMID: 22975761. Pubmed Central PMCID: PMC3743447
13. Wishart GC, Bajdik CD, Dicks E, et al. PREDICT Plus: development and validation of a prognostic model for early breast cancer that includes HER2. *Br J Cancer.* 2012;107(5):800–7. PubMed PMID: 22850554. Pubmed Central PMCID: PMC3425970
14. Ravdin PM. A computer program to assist in making breast cancer adjuvant therapy decisions. *Semin Oncol.* 1996;23(1 Suppl 2):43–50. PubMed PMID: 8614844
15. Ravdin PM, Siminoff LA, Davis GJ, et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol.* 2001;19(4):980–91. PubMed PMID: 11181660
16. Warren JL, Harlan LC, Fahey A, et al. Utility of the SEER-Medicare data to identify chemotherapy use. *Med Care.* 2002;40(8 Suppl):IV-55-61. PubMed PMID: 12187169
17. Warren JL, Klabunde CN, Schrag D, et al. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. *Med Care.* 2002;40(8 Suppl):IV-3-18. PubMed PMID: 12187163
18. Hosseini H, Obradovic MM, Hoffmann M, et al. Early dissemination seeds metastasis in breast cancer. *Nature.* 2016. PubMed PMID: 27974799.
19. Cardoso F, van't Veer LJ, Bogaerts J, et al. 70-Gene signature as an aid to treatment decisions in early-stage breast cancer. *N Engl J Med.* 2016;375(8):717–29. PubMed PMID: 27557300
20. Campbell HE, Taylor MA, Harris AL, Gray AM. An investigation into the performance of the Adjuvant! Online prognostic programme in early breast cancer for a cohort of patients in the United Kingdom. *Br J Cancer.* 2009;101(7):1074–84. PubMed PMID: 19724274. Pubmed Central PMCID: PMC2768087
21. de Glas NA, van de Water W, Engelhardt EG, et al. Validity of Adjuvant! Online program in older patients with breast cancer: a population-based study. *Lancet Oncol.* 2014;15(7):722–9. PubMed PMID: 24836274
22. Mook S, Schmidt MK, Rutgers EJ, et al. Calibration and discriminatory accuracy of prognosis calculation for breast cancer with the online Adjuvant! program: a hospital-based retrospective cohort study. *Lancet Oncol.* 2009;10(11):1070–6. PubMed PMID: 19801202
23. Loprinzi CL, Thome SD. Understanding the utility of adjuvant systemic therapy for primary breast cancer. *J Clin Oncol.* 2001;19(4):972–9. PubMed PMID: 11181659
24. Shachar SS, Muss HB. Internet tools to enhance breast cancer care. *NPJ Breast Cancer.* 2016;2:16011.
25. Siminoff LA, Gordon NH, Silverman P, et al. A decision aid to assist in adjuvant therapy choices for breast cancer. *Psychooncology.* 2006;15(11):1001–13. PubMed PMID: 16511899
26. Peele PB, Siminoff LA, Xu Y, Ravdin PM. Decreased use of adjuvant breast cancer therapy in a randomized controlled trial of a decision aid with individualized risk information. *Med Decis Mak.* 2005;25(3):301–7. PubMed PMID: 15951457
27. Joishy SK, Driscoll JC. The ailments of cancer registries: a proposal for remedial education. *J Cancer Educ.* 1989;4(1):17–31. PubMed PMID: 2641505
28. Brewster D, Crichton J, Muir C. How accurate are Scottish cancer registration data? *Br J Cancer.* 1994;70(5):954–9. PubMed PMID: 7947104. Pubmed Central PMCID: PMC2033548
29. Brewster D. Improving the quality of cancer registration data. *J R Soc Med.* 1995;88(5):268–71. PubMed PMID: 7636820. Pubmed Central PMCID: PMC1295197
30. Wong HS, Subramaniam S, Alias Z, et al. The predictive accuracy of PREDICT: a personalized decision-making tool for Southeast Asian women with breast cancer. *Medicine (Baltimore).* 2015;94(8):e593. PubMed PMID: 25715267. Pubmed Central PMCID: PMC4554151
31. Harris LN, Ismaila N, McShane LM, et al. Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer: American society of clinical oncology clinical practice guideline. *J Clin Oncol.* 2016;34(10):1134–50. PubMed PMID: 26858339. Pubmed Central PMCID: PMC4933134
32. Gene expression profiling and expanded immunohistochemistry tests for guiding adjuvant chemotherapy decisions in early breast cancer management: MammaPrint, Oncotype DX, IHC4 and Mammostrat. London, UK: National Institute for Health and Care Excellence (NICE), 2013 Contract No.: Diagnostics Guidance 10.
33. Buus R, Sestak I, Kronenwett R, et al. Comparison of EndoPredict and EPclin with oncotype DX recurrence score for prediction of risk of distant recurrence after endocrine therapy. *J Natl Cancer Inst.* 2016;108(11). PubMed PMID: 27400969. Pubmed Central PMCID: PMC5241904.

34. Harowicz MR, Robinson TJ, Dinan MA, et al. Algorithms for prediction of the Oncotype DX recurrence score using clinicopathologic data: a review and comparison using an independent dataset. *Breast Cancer Res Treat.* 2017. PubMed PMID: 28064383.
35. Tsai M. L. US, Treece T., Lo S. S., PROMIS Investigator Group. The 70-gene signature to provide risk stratification and treatment guidance for patients classified as intermediate by the 21-gene assay. *J Clin Oncol.* 2016 ASCO Annual Meeting 2016.
36. Greaves M. Evolutionary determinants of cancer. *Cancer Discov.* 2015;5(8):806–20. PubMed PMID: 26193902. Pubmed Central PMCID: PMC4539576

Submit your next manuscript to BioMed Central
and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

