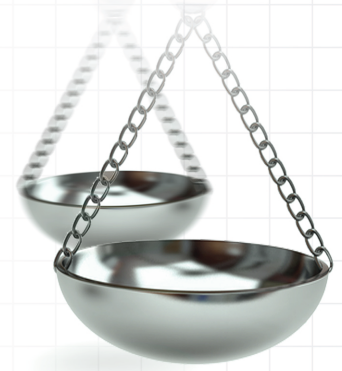




Letter to the Editor

Malek B Hannouf¹, Dale Muzzey², Ralf Kronenwett³ & Johnathan M Lancaster^{*,2}

¹Department of Surgery, Schulich School of Medicine & Dentistry, Western University, London, ON, Canada, N6A 3K7

²Myriad Genetics, Inc., Salt Lake City, UT 84108, USA

³Myriad International GmbH, Cologne, Germany

*Author for correspondence: Tel.: +1 801 505 5090; jmlgsdad@gmail.com

Journal of **Comparative
Effectiveness Research**

First draft submitted: 15 April 2019; Accepted for publication: 12 June 2019; Published online: 19 November 2019

Keywords: breast cancer • chemotherapy de-escalation • comparative effectiveness research • economics • evidence-based practice • meta-analysis • multi-gene assays

We read with interest the recent article by Hochheiser *et al.* regarding the economic impact of four multi-gene assays (MGAs) used to make decisions regarding chemotherapy use in women with newly diagnosed estrogen receptor-positive, HER2-negative breast cancer: 21-gene recurrence score (Oncotype DX[®]), 70-gene signature (MammaPrint[®]), PAM50 (commercial; Prosigna[®]) and 12-gene molecular score (EndoPredict[®]) [1]. We wish to highlight several deficiencies in the study design and the resultant conclusions.

The authors described the variables that were utilized in their cost–effectiveness model, which included two elements related to the risk of distant recurrence according to MGA risk stratification: the 10-year probability of recurrence without chemotherapy (i.e., prognostic performance) and the 10-year relative risk of recurrence with and without chemotherapy (i.e., predictive performance). These values were then applied to decision-impact data (i.e., how testing with each MGA impacted chemotherapy treatment decisions) to model the cost–effectiveness of each MGA. Hochheiser *et al.* [1] followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines to compile all available and applicable decision-impact literature for the four MGAs. However, the authors did not use this same approach to compile literature reporting on the ability of the MGAs to stratify the 10-year risk of distant recurrence.

There are several significant issues regarding the values for the prognostic performance of the MGAs that the authors used as inputs to their model. Hochheiser *et al.* [1] cite a meta-analysis from the Early Breast Cancer Trialists' Collaborative Group [2] as the source of the input prognostic values for the 12-gene molecular score and PAM50 (commercial) score. The authors state that this meta-analysis “*provided commonly cited rates for distant recurrence*”; however, this citation predates the MGAs and does not contain any molecular data. Therefore, it is not possible to use this meta-analysis to obtain input values for the probability of recurrence without chemotherapy according to MGA risk group for the 12-gene or PAM50 (commercial) scores if the MGAs themselves did not exist yet. As such, the base prognostic variables used for these MGAs are questionable and may be invalid and, by association, so are any conclusions drawn from them.

The EndoPredict test result is a combined clinico-molecular score (EPclin) that incorporates tumor size and nodal status with a 12-gene molecular score. Hochheiser *et al.* [1], used decision-impact data for the EPclin score in their model; however, the base case variables used for the prognostic and predictive performance appear to be for the 12-gene molecular score and not the EPclin score. This is inappropriate for two reasons. First, there is inconsistency within the authors' model regarding the use of the 12-gene molecular score and EPclin score. Second, the EPclin score is the actual test result provided to clinicians for use in treatment decisions and would have been the appropriate result to model for this study.

Notably, there is a wealth of published literature that includes prognostic data for both the EPclin score [3–6] and PAM50 (commercial) score [3,7–9]. It is not obvious why the authors did not utilize the available data in their model, as the published data clearly demonstrate that the input values used in the cost–effectiveness model are wrong. For instance, a recent report from Sestak *et al.* [3], which predated the publication of Hochheiser *et al.* [1], by 9 months, showed that the 10-year risk of distant recurrence for women treated with endocrine therapy alone who had a high-risk EPclin score was 22.1%. This is substantially higher than the 15% risk modeled by Hochheiser

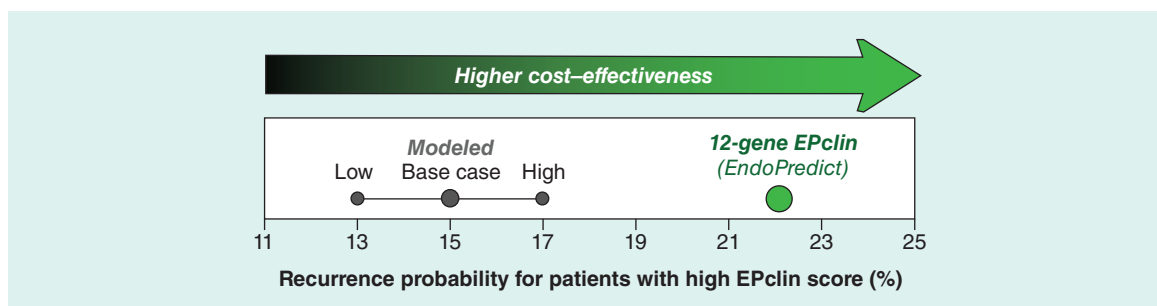


Figure 1. Comparison of the average ('base case') and range of prognostic performance of the 12-gene multi-gene assay modeled in the Hochheiser *et al.* cost-effectiveness analysis (gray) when compared with the value published by Sestak *et al.* (green).

et al. [1] (Figure 1). This directly demonstrates that the input prognostic values grossly underestimate the true prognostic performance of the 12-gene MGA (EPclin score) and, consequently, that the authors' assumptions would underestimate the cost-effectiveness of EndoPredict.

Beyond providing a citation, the authors do not provide a detailed description about how the prognostic input values for the four MGAs were obtained (Hochheiser *et al.* [1], Table 1). Strangely, the input prognostic values and ranges given for the low-risk categories of all four MGAs are identical, despite the fact that three different citations are referenced. The input prognostic values for the high-risk 21-gene recurrence score and PAM50 (commercial) categories are also identical. However, the study by Sestak *et al.* [3] included a direct comparison of the prognostic performance of the 21-gene recurrence score, PAM50 (commercial), and the 12-gene MGA (EPclin), showing that the prognostic performance of these MGAs differed from one another [3]. In light of the Sestak *et al.* data and without a clear explanation of how the prognostic input values were obtained in Hochheiser *et al.* [1] the assumed prognostic performance of the MGAs in the latter's cost-effectiveness analysis are not appropriate. The lack of clear provenance for prognostic input values is especially troubling in light of the authors' own sensitivity analyses, which showed that the prognostic performance of the MGAs (probability of recurrence without chemotherapy) is the most significant contributor to the outcome of their cost-effectiveness analysis (Hochheiser *et al.* [1], Supplementary Figure 2). Despite this fact, there is limited information included in the manuscript to inform the reader that the prognostic performance of each test has the greatest impact on the overall conclusions.

The authors also failed to provide sufficient information regarding how the predictive input values were obtained from the cited literature. For the 21-gene recurrence score, the authors simply state that "*the relative reduction of distant recurrence with chemotherapy by 21-gene Recurrence Score risk group was based on Paik et al.*" However, the range of relative risk of recurrence (Hochheiser *et al.*, Table 1) differs from what was reported in the cited paper (Paik *et al.* [10] Figure 3A) [10]. For high-risk patients, where Paik *et al.* reported a mean value of 0.74, the risk range utilized by Hochheiser *et al.* (0.69–0.78) does not match the 95% CIs observed in Paik *et al.* [10] (0.47–0.87). The assumption of a narrower and less skewed range than what was empirically observed in Paik *et al.* likely has substantial ramifications on the modeled cost-effectiveness of the 21-gene recurrence score, as a narrower and less downward-skewed range in high-risk patients could spuriously increase confidence in the claim of cost-effectiveness. This brings into question the omission of detailed information regarding study methodology or rationale.

In addition to the range of predictive power being questionable, the average value is questionable as well; others have shown that the analysis by Paik *et al.* [10] was flawed due to an overlap between the training and validation cohorts, resulting in an overestimation of the predictive power of the 21-gene score [11]. When a more appropriate analysis was performed by Ioannidis *et al.* [11], the 21-gene score did not predict benefits from chemotherapy [11].

In the absence of available literature providing the 10-year relative risk of recurrence for the 12-gene and PAM50 (commercial) scores, the authors should not have performed a comparison using this variable, or at the very least, they should have indicated their uncertainty with wide error boundaries. The incorrect input values used for both the prognostic and predictive values of the MGAs overestimate the cost-effectiveness of the 21-gene recurrence score and underestimate the performance of the other MGAs. As a result, the authors' findings that only the 21-gene recurrence score resulted in a decrease in healthcare costs was a foregone conclusion.

We agree with the authors that an understanding of the relative cost–effectiveness of available MGAs is relevant for clinical use. An appropriate cost–effectiveness comparison could have been performed utilizing prognostic data, which is available for all four MGAs. This approach has been utilized in several cost–effectiveness models, which show that testing with the PAM50 (commercial) or the 12-gene MGA result in a net reduction in healthcare costs [12–16]. As the cost–effectiveness of breast cancer MGAs is an important part of their clinical utilization, accurate reporting is critical, and, therefore, robust study design is imperative to produce informative results.

Financial & competing interests disclosure

D Muzzey, R Kronenwett and JM Lancaster were employed by Myriad Genetics at the time of this work. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

References

1. Hochheiser L, Hornberger J, Turner M, Lyman GH. Multi-gene assays: effect on chemotherapy use, toxicity and cost in estrogen receptor-positive early stage breast cancer. *J. Comp. Eff. Res.* 8(5), 289–304 (2019).
2. Early Breast Cancer Trialists' Collaborative Group. *Treatment of Early Breast Cancer: A Systematic Overview of all Available Randomized Trials of Adjuvant Endocrine and Cytotoxic Therapy Volume 1: Worldwide Evidence 1985–1990* Oxford University Press, Oxford, UK, 1985–1990.
3. Sestak I, Buus R, Cuzick J *et al.* Comparison of the performance of 6 prognostic signatures for estrogen receptor-positive breast cancer: a secondary analysis of a randomized clinical trial. *JAMA Oncol.* 4(4), 545–553 (2018).
4. Filipits M, Rudas M, Jakesz R *et al.* A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. *Clin. Cancer Res.* 17(18), 6012 (2011).
5. Dubsy 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. *Ann. Oncol.* 24(3), 640–647 (2012).
6. Dubsy P, Brase JC, Jakesz R *et al.* The EndoPredict score provides prognostic information on late distant metastases in ER+/HER2-breast cancer patients. *Brit. J. Cancer* 109, 2959–2964 (2013).
7. 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.* 31, 2783–2790 (2013).
8. Filipits M, Nielsen TO, Rudas M *et al.* The PAM50 risk-of-recurrence score predicts risk for late distant recurrence after endocrine therapy in postmenopausal women with endocrine-responsive early breast cancer. *Clin. Cancer Res.* 20(5), 1298–1305 (2014).
9. Gnani 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.* 25(2), 339–345 (2014).
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(23), 3726–3734 (2006).
11. Ioannidis JP. Is molecular profiling ready for use in clinical decision making? *Oncologist* 12(3), 301–311 (2007).
12. Hannouf MB, Zaric GS, Blanchette P *et al.* Cost–effectiveness analysis of multigene expression profiling assays to guide adjuvant therapy decisions in women with invasive early stage breast cancer. *Pharmacogenomics J.* doi:10.1038/s41397-019-0089-x (2019).
13. Wang SY, Chen T, Dang W, Mougalian SS, Evans SB, Gross CP. Incorporating tumor characteristics to maximize 21-gene assay utility: a cost–effectiveness analysis. *J. Natl Compr. Canc. Netw.* 17(1), 39–46 (2019).
14. Blank PR, Filipits M, Dubsy P *et al.* Cost–effectiveness analysis of prognostic gene expression signature-based stratification of early breast cancer patients. *Pharmacoeconomics* 33(2), 179–190 (2015).
15. Wang SY, Dang W, Richman I, Mougalian SS, Evans SB, Gross CP. Cost–effectiveness analyses of the 21-gene assay in breast cancer: systematic review and critical appraisal. *J. Clin. Oncol.* 36(16), 1619–1627 (2018).
16. Hinde S, Theriou C, May S *et al.* The cost–effectiveness of EndoPredict to inform adjuvant chemotherapy decisions in early breast cancer. *Health Pol. Tech.* 8(1), 75–83 (2019).