

Evaluation of three commercial progesterone receptor assays in a single tamoxifen-treated breast cancer cohort

Elizabeth N Kornaga¹, Alexander C Klimowicz², Natalia Guggisberg¹, Travis Ogilvie³, Don G Morris^{1,4}, Marc Webster⁴ and Anthony M Magliocco⁵

¹Translational Laboratories, Tom Baker Cancer Centre, Alberta Health Services, Calgary, Alberta, Canada;

²Boehringer Ingelheim Pharmaceuticals, Ridgefield, CT, USA; ³Department of Pathology, Calgary Laboratory Services, Calgary, Alberta, Canada; ⁴Department of Oncology, Tom Baker Cancer Centre and University of Calgary, Calgary, Alberta, Canada and ⁵Department of Anatomic Pathology, H Lee Moffitt Cancer Center and Research Institute, Tampa, FL, USA

Estrogen receptor and progesterone receptor status are routinely assessed using immunohistochemistry assays to assist in patient prognosis and clinical management. Three commonly utilized autostainer vendors—Dako, Leica and Ventana—provide ready-to-use progesterone receptor assays; however, they have never been directly compared in a single breast cancer cohort. We looked at three immunohistochemical progesterone receptor assays, in addition to original ligand-binding assay results, in a single retrospective, tamoxifen-treated breast cancer cohort to investigate inter- and intra-observer agreement, concordance, prognostic ability and measures of test performance. All immunohistochemical assays utilized the manufacturer's specified protocols. Five-year disease-free survival was the endpoint of interest, and multivariate models were adjusted for lymph node status, tumor grade, tumor size and human epidermal growth factor 2 status. All assays showed substantial to almost perfect agreement between the three observers (Dako $\kappa = 0.69$ – 0.90 ; Leica $\kappa = 0.70$ – 0.89 ; and Ventana $\kappa = 0.78$ – 0.94) and concordance (Dako/Leica $\kappa = 0.81$; Dako/Ventana $\kappa = 0.78$; and Leica/Ventana $\kappa = 0.82$). Univariate survival analyses showed that only the ligand-binding assay, Dako and Ventana assays achieved statistical significance. No statistically significant results were seen in multivariate models, although a strong trend was seen with the Ventana progesterone receptor assay. All assays performed similarly with regards to measures of test performance with ligand-binding assay set as the reference, and all immunohistochemical assays outperformed the ligand-binding assay in regards to 5-year disease-free survival. Despite similar agreement and concordance with the progesterone receptor assays, clear differences were noted with regards to 5-year disease-free survival. Additional survival analyses suggest that clinical utility of estrogen receptor assays vary when investigated in combination with progesterone receptor.

Modern Pathology (2016) 29, 1492–1500; doi:10.1038/modpathol.2016.151; published online 26 August 2016

Endocrine therapy for hormone receptor-positive breast cancers is a standard treatment that is proven to improve patient outcome.^{1–3} Currently, both estrogen receptor and progesterone receptors are routinely analyzed using immunohistochemical techniques on formalin-fixed paraffin-embedded specimens to determine patient prognosis and management.^{1–4} Patient management decisions are

predominantly made on the results of the estrogen receptor assay; the role of progesterone receptor is not clearly established according to ASCO/CAP guidelines⁴ despite many groups showing the prognostic and predictive value of progesterone receptor as a breast cancer biomarker.^{5–19}

Immunohistochemical assays slowly replaced the ligand-binding assay for hormone receptor determination between the mid-90s and early 2000s. Following the adoption of immunohistochemical techniques was the development of automated immunostainers and adoption of multiple different progesterone receptor antibody clones. Most early assays were laboratory developed, and although some laboratories still use in-house developed

Correspondence: Dr AM Magliocco, MD, FRCPC, FCAP, Department of Anatomic Pathology, H Lee Moffitt Cancer Center and Research Institute, 12902 Magnolia Drive, Tampa, FL 33612, USA. E-mail: Anthony.Magliocco@moffitt.org

Received 16 May 2016; revised 11 July 2016; accepted 14 July 2016; published online 26 August 2016

assays, many clinical laboratories have opted to use platform-specific, ready-to-use assays developed by the major manufacturers of automated immunohistochemistry stainers: DakoCytomation (Dako), Leica Microsystems (Leica) and Ventana Medical Systems (Ventana). Ready-to-use assays minimize variability, and improve reproducibility and consistency.^{20–28} We have previously compared these platform-specific ready-to-use estrogen receptor assays,²⁹ but a direct comparison of the progesterone receptor assays on a single clinical outcome series has never been performed.

Here we present a systematic comparison of three vendor-specific progesterone receptor ready-to-use immunohistochemical assays using a retrospective, tamoxifen-treated breast cancer cohort to: (1) assess observer agreement and concordance; (2) explore the prognostic value of the progesterone receptor ligand-binding assay and immunohistochemical ready-to-use assays; and (3) explore the standard measures of test performance using (a) ligand-binding, or (b) 5-year disease-free survival, as the reference standards.

Materials and methods

Patient Cohort

The Calgary Tamoxifen Breast Cancer Cohort (Calgary cohort) is a retrospective database containing demographic, clinical and pathological data for breast cancer patients diagnosed between 1985 and 2000 at the Tom Baker Cancer Centre in Calgary, Alberta, Canada and has been previously described.²⁹ Briefly, a total of 532 primary tumors were identified that received surgical intervention, adjuvant tamoxifen endocrine therapy without primary or adjuvant chemotherapy, had no prior cancer diagnosis (except non-melanoma skin cancer) and had available archival formalin-fixed paraffin-embedded tissue with confirmed invasive cancer. Archival tissue was reviewed, and pathologist confirmed invasive tumor tissue was placed into tissue microarrays.³⁰ Adjuvant tamoxifen treatment was offered regardless of hormone receptor status in this cohort as there was no definitive evidence at the time these patients were diagnosed that receptor-negative patients did not respond to tamoxifen.

Clinical Assessment of Hormone Receptor Status

At the time of diagnosis for patients in this cohort (1985–2000), hormone status was predominantly assessed using dextran-coated charcoal (ligand)-binding assay.³¹ Tumors were considered to be receptor positive if the ligand-binding assay results were >10 fmol/ml. Cases that did not have ligand-binding data available were assessed by early immunohistochemical methods. Human epidermal growth factor 2 (HER2) was not assessed clinically during this time period, and was retrospectively

assessed using the Dako HercepTest pharmDx kit, as previously described.²⁹

Immunohistochemistry

Detailed information regarding immunohistochemical staining for estrogen and progesterone receptor has been previously described for this cohort.²⁹ Briefly, all RTU assays were performed on corresponding vendor-specific autostainer systems according to the manufacturer's instructions. Cell line controls from the Dako ER/PR pharmDx kit and HercepTest pharmDx kit were run in addition to a laboratory built reference tissue microarrays. All tissue microarrays were reviewed and showed consistent staining, suggesting that the antibody was uniformly applied during the staining process.

Immunohistochemistry Scoring

HER2, estrogen and progesterone receptor were manually scored following the ASCO/CAP guidelines.^{4,32} Specifically, HER2 was scored as either 0, 1+, 2+ or 3+, and tumors were considered positive if the HER2 average of the replicate cores for each case were >2. Estrogen and progesterone receptor status was assessed following the Allred scoring method,³³ and tumors were considered positive if they had an overall Allred score of 3 or higher in at least one of the replicate cores. Scoring was performed as a consensus between two highly trained researchers (observer 1), and two expert pathologists (observers 2 and 3). Slides were rescored by observers 1 and 3 three months after initial review to assess intra-observer reproducibility.

Statistical Analysis

All statistical analyses were performed using Stata 12 (StataCorp LP). The kappa statistic was used to measure inter- and intra-observer agreement, and inter-platform comparisons.³⁴ The event under study was 5-year disease-free survival, defined as time from diagnosis to local recurrence, metastatic disease or death from breast cancer. Kaplan–Meier curves were analyzed using the log-rank test at 5-year disease-free survival and Cox proportional hazard regression was performed to estimate hazard ratios, adjusting for lymph node status, tumor grade, tumor size and HER2 status. All survival analyses were performed in estrogen receptor-positive cases, as determined by the corresponding vendor estrogen receptor assay results. Subjects were excluded from the multivariate models if there were missing data for any of these variables. The proportional hazard assumptions were tested by assessing log–log survival curves as well as the goodness-of-fit using the Schoenfeld residuals test. The ligand-binding assay

results were used as the reference standard for calculating measures of test performance, and additional calculations were performed, comparing ligand-binding and the immunohistochemical assays, with 5-year disease-free survival as the reference standard.³⁵

Results

Agreement

Inter-observer agreement was evaluated for all three platforms by comparing observer interpretation of each core for progesterone receptor staining between three observers, and was measured using the kappa statistic. Inter-observer interpretations on the Ventana platform consistently showed the strongest agreement for progesterone receptor, with $\kappa=0.94$ between observers 1 and 2, $\kappa=0.78$ between observers 1 and 3, and $\kappa=0.84$ between observers 2 and 3. Inter-observer agreement for progesterone receptor interpretation on the Leica platform showed substantial to almost perfect agreement with $\kappa=0.89$ between observers 1 and 2, $\kappa=0.70$ between observers 1 and 3, and $\kappa=0.79$ between observers 2 and 3. Kappa values for interpretation with the Dako platform also showed substantial to almost perfect agreement, with $\kappa=0.90$ between observers 1 and 2, $\kappa=0.69$ between observers 1 and 3, and $\kappa=0.79$ between observers 2 and 3 for progesterone receptor.

The stained slides were rescored 3 months after initial review by observers 1 and 3, and intra-observer agreement was calculated. For observer 1, all three platforms had almost perfect agreement (Dako and Ventana, $\kappa=0.98$; Leica, $\kappa=0.94$). For observer 3, all platforms also had almost perfect agreement (Dako, $\kappa=0.93$; Leica, $\kappa=0.92$; Ventana, $\kappa=0.84$).

Concordance

Progesterone receptor status for the three ready-to-use assays was compared as a core by core analysis, and agreement was again assessed using the kappa statistic. Substantial agreement for progesterone receptor was seen between the cores stained with the Dako and Ventana assays, $\kappa=0.78$ with 49 discordant cases ($n=23$, Dako negative/Ventana positive; $n=26$ Dako positive/Ventana negative). Substantial agreement for progesterone receptor was also seen for Dako and Leica assays, $\kappa=0.81$ with 47 discordant cases ($n=19$ Dako negative/Leica positive; $n=28$ Dako positive/Leica negative). Agreement between Leica and Ventana was similar, $\kappa=0.82$, with 44 discordant cases ($n=23$ Leica negative/Ventana positive; $n=21$ Leica positive/Ventana negative). Example images of discordance with progesterone receptor staining are presented in Figure 1.

Univariate Analysis of Progesterone Receptor

Cases were dichotomized into positive and negative as previously described,^{4,31} and Kaplan–Meier survival curves for 5-year disease-free survival were analyzed for progesterone receptor in the estrogen receptor-positive cases (Figure 2). The log-rank test was used to compare positive and negative groups and hazard ratios were calculated to compare relative survival. The ligand binding, as well as the Dako and Ventana assays, achieved significance with the log-rank test (ligand-binding assay Figure 2a, $P=0.049$; Dako Figure 2b, $P=0.043$; Ventana Figure 2d; $P=0.033$), whereas the Leica assay failed to show significance (Figure 2c, $P=0.359$). Univariate proportional hazard analysis showed significance for the Dako and Ventana assays ($P=0.046$ and $P=0.036$, respectively), and near significance for the ligand-binding assay ($P=0.051$); the Leica assay failed to reach significance ($P=0.361$). The corresponding hazard ratios and 95% confidence intervals for all assays are presented in Table 1.

Multivariate Analysis of Progesterone Receptor

Cox proportional hazard models were analyzed for the effect of progesterone receptor, as determined by the ligand-binding and immunohistochemical assays for each platform, and are presented in Table 2. All analyses were adjusted for lymph node status, tumor grade, tumor size and HER2 status. None of the assays achieved significance in the multivariate models (Table 2); however, the Ventana assay did show a trend toward significance ($P=0.090$).

Measures of Test Performance

Sensitivity, specificity, positive predictive value, negative predictive value and accuracy were calculated for all platforms, using the ligand-binding assay results as the reference standard (Table 3). All ready-to-use assays performed comparably to the ligand-binding assay. Remarkably sensitivity was strong for all assays ($>96\%$), and positive predictive value, negative predictive value and accuracy were similar, between 72 and 77%, for all immunohistochemical assays, whereas specificity was consistently low (23–25%).

Measures of test performance were also calculated for the ligand-binding and ready-to-use immunohistochemical assays with 5-year disease-free survival used as the reference standard (Table 4). All three immunohistochemical assays performed similar to the ligand-binding assay for sensitivity, specificity, positive predictive value and accuracy. Interestingly, all measures of test performance for the immunohistochemical assays greatly improved, and moreover, outperformed the ligand-binding assay in respect to sensitivity (92–93% vs 73%) and accuracy (77–79% vs 68%). Positive predictive and negative

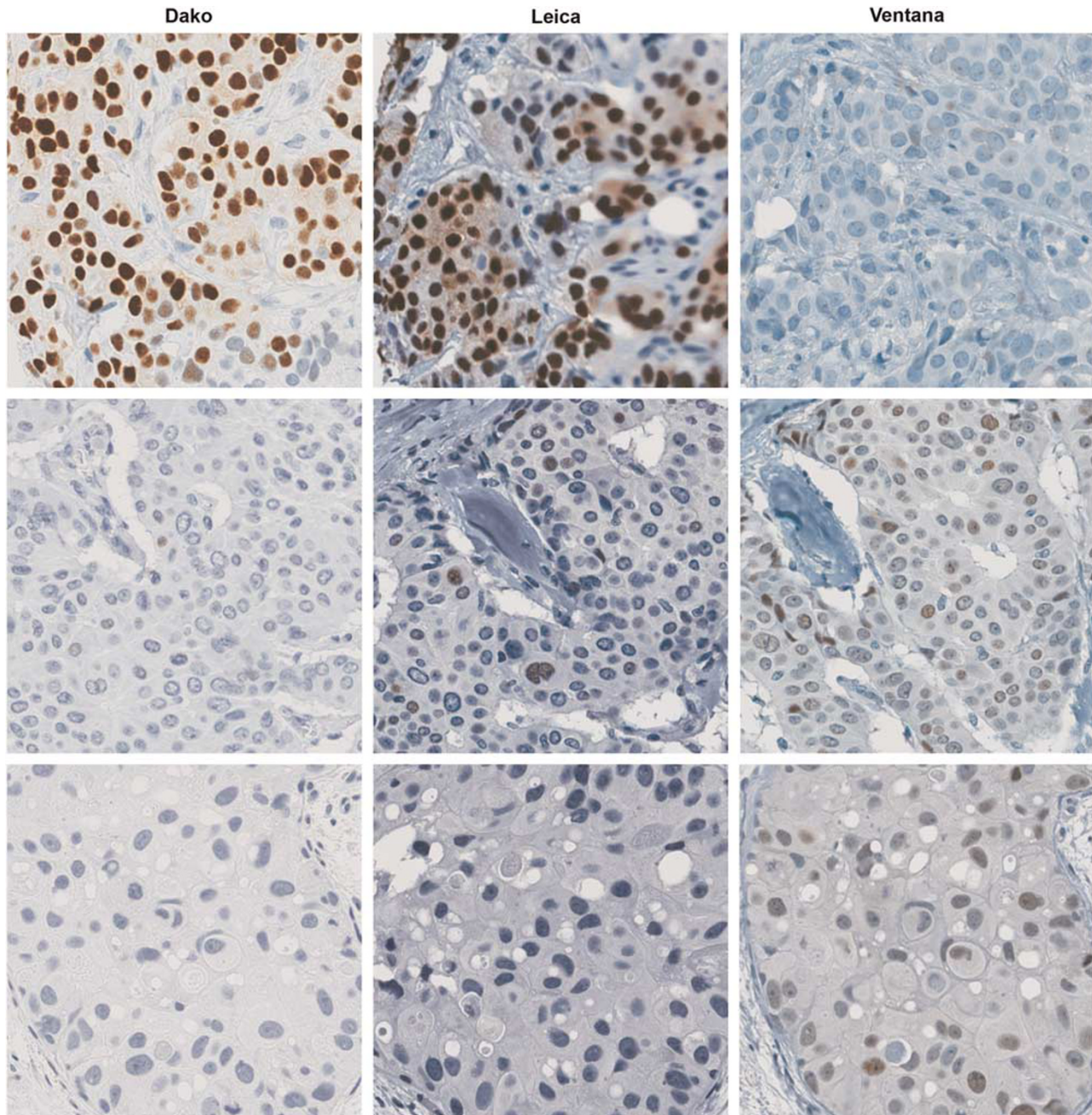


Figure 1 Example images of progesterone receptor assay discordance between vendor ready-to-use assays for Dako (clone PgR 1294), Leica (clones 16 and SAN27) and Ventana (clone 1E2).

predictive values were similar across all assays, whereas the ligand-binding assay showed the highest specificity (39% vs 11–15% for the immunohistochemical assays).

Discussion

This study presents a systematic comparison of three ready-to-use progesterone receptor immunohistochemical assays using a retrospective, tamoxifen-treated, breast cancer cohort. Utilizing a ready-to-use

assay along with an automated staining platform for immunohistochemical analysis of progesterone receptor offers increased reproducibility and standardization, minimizing potential analytical errors that may lead to incorrect test results.^{20,24,36} We have previously reported on our comparison of the ready-to-use estrogen receptor assays with this cohort,²⁹ and many of the points discussed are valid with the progesterone receptor evaluation presented here. Specifically, ready-to-use antibodies are titrated by the vendor to ensure optimal and consistent results, negating the need for laboratory personnel to

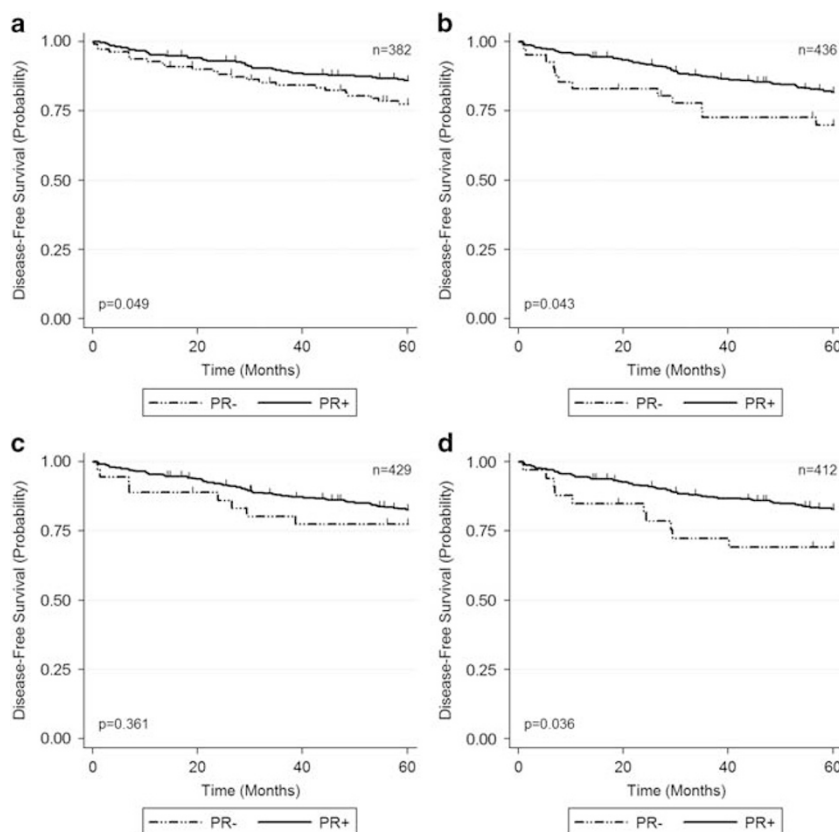


Figure 2 Univariate analysis, including log-rank *P*-values, of progesterone receptor negative (PR-) vs positive (PR+) in estrogen receptor positive cases: (a) ligand-binding assay (LBA); (b) Dako; (c) Leica; and (d) Ventana. Total sample size for univariate analysis of each assay is indicated in the upper right corner.

perform rigorous titrations, and minimizing potential inter- and intra-laboratory variability (lot to lot evaluations should still be performed). In addition, the improved consistency provided by these assays allows for increased reproducibility in reporting the results between observers. Taken together, this leads to improved performance and reliability, which is particularly important for a clinically utilized assay.

We evaluated platform concordance by a core-to-core comparison, and noted that concordance was similar between platforms ($\kappa=0.78$ – 0.82). Similarly, all ready-to-use assays had similar rates of progesterone receptor negativity (Dako 12%; Leica 13%; Ventana 11%). As we previously noted in our estrogen receptor comparison, counterstaining varied between platforms (Figure 1). As we followed the manufacturer's recommended protocols, we did not augment this step to provide counterstain consistency across the platforms, and this may have contributed to some of the discordance seen. In addition, all vendors utilized different antibodies for their ready-to-use assays, each potentially recognizing different epitopes. Specifically, Dako utilized antibody clone progesterone receptor 1294 (recognizes A and B isoforms, binding to N-terminal amino acids 165–534), Leica was a progesterone receptor cocktail including clone 16 (recognizes A isoform,

Table 1 Univariate Cox proportional hazard models for the LBA and Dako, Leica and Ventana ready-to-use immunohistochemical assays

	HR	95% Confidence interval	P-value
Ligand binding	0.60	(0.36–1.00)	0.051
Dako	0.54	(0.29–0.99)	0.046
Leica	0.71	(0.34–1.48)	0.361
Ventana	0.49	(0.25–0.96)	0.036

Abbreviation: LBA, ligand-binding assay.

binding to N-terminal, exact epitope unknown) and clone SAN27 (recognizes B isoform, binding to 164 amino-acid sequence in N-terminal unique to B isoform), whereas Ventana make use of the progesterone receptor antibody 1E2 clone (recognizes A and B isoforms, specific binding unknown), which is the more likely cause of discordance.^{37,38}

Our study utilized archival tissue specimens, and pre-analytical variables could not be controlled and thus, may be another potential source for some of the observed discordance. It has previously been shown that pre-analytical variation can negatively affect immunohistochemical biomarkers, specifically

Table 2 Multivariate Cox proportional hazard models for the LBA and Dako, Leica and Ventana ready-to-use immunohistochemical assays

	HR	95% Confidence interval	P-value
Ligand binding	0.75	(0.39–1.43)	0.377
Dako	0.65	(0.31–1.37)	0.261
Leica	0.83	(0.33–2.09)	0.693
Ventana	0.5	(0.22–1.12)	0.090

Abbreviation: LBA, ligand-binding assay.

All models were adjusted for lymph node status, tumor grade, tumor size and HER2 status.

Table 3 Measures of test performance comparing progesterone receptor ready-to-use immunohistochemical assays to ligand-binding assay results

	Dako	Leica	Ventana
<i>Sensitivity (%)</i>	96.2%	97.0%	97.4%
TP	227	230	226
TP + FN	236	237	232
<i>Specificity (%)</i>	22.8%	25.0%	22.5%
TN	23	24	20
TN + FP	101	96	89
<i>Positive predictive value (%)</i>	74.4%	76.2%	76.6%
TP	227	230	226
TP + FP	305	302	295
<i>Negative predictive value (%)</i>	71.9%	77.4%	76.9%
TN	23	24	20
TN + FN	32	31	26
<i>Accuracy (%)</i>	74.2%	76.3%	76.6%
TP + TN	250	254	246
TP + FP + TN + FN	337	333	321

Abbreviations: FN, false negative; FP, false positive; TN, true negative; TP, true positive.

Table 4 Measure of test performance comparing LBA and ready-to-use immunohistochemical assays to 5-year disease-free survival

	LBA	Dako	Leica	Ventana
<i>Sensitivity (%)</i>	73.1%	91.8%	92.1%	93.2%
TP	234	324	326	315
TP + FN	320	353	354	338
<i>Specificity (%)</i>	38.7%	14.5%	10.7%	13.5%
TN	24	12	8	10
TN + FP	62	83	75	74
<i>Positive predictive value (%)</i>	86.0%	82.0%	83.0%	83.1%
TP	234	324	326	315
TP + FP	272	395	393	379
<i>Negative predictive value (%)</i>	21.8%	29.3%	22.2%	30.3%
TN	24	12	8	10
TN + FN	110	41	36	33
<i>Accuracy (%)</i>	67.5%	77.1%	77.9%	78.9%
TP + TN	258	336	334	325
TP + FP + TN + FN	382	436	429	412

Abbreviations: FN, false negative; FP, false positive; LBA, ligand-binding assay; TN, true negative; TP, true positive.

in breast cancer;^{25,26,39} however, a recent publication has shown that the effects of delayed fixation is consistent among the progesterone receptor clones evaluated in this study,³⁸ suggesting that all clones would be equally affected by unknown pre-analytical variables and an unlikely source for the discordance observed in our study.

The prognostic and/or predictive significance of progesterone receptor in breast cancer patients is still debated;^{4,40} however, there is sufficient evidence to suggest it has prognostic value^{5–19} and we subsequently investigated all progesterone receptor assays with regards to patient survival in estrogen receptor-positive cases, as determined by the corresponding vendor-specific estrogen receptor assay result. This exploration yielded interesting variations between the assays. Univariate analysis of the ligand-binding, Dako and Ventana assays all showed significance, or near significance (Figure 2 and Table 1), whereas the Leica assay did not show significance (HR=0.71 (95% confidence interval: 0.34–1.48), $P=0.361$). Even more intriguing, only the Ventana progesterone receptor assay neared significance in the multivariate Cox model that adjusted for lymph node status, tumor grade, tumor size and HER2 status (HR=0.50 (95% CI: 0.22–1.12), $P=0.090$) for 5-year disease-free survival. Unlike our evaluation of the different estrogen receptor assays where we reported functional equivalency with the univariate and multivariate models,²⁹ these results suggest the Ventana progesterone receptor assay utilizing the 1E2 clone may be superior to the other assays and clones investigated. These findings may partially, or wholly, explain discordance in the role of progesterone receptor in the literature.

We also assessed standard measures of test performance—sensitivity, specificity, positive predictive value, negative predictive value and accuracy. All assays performed similarly in all measures when the ligand-binding assay was set as the reference standard (Table 3). Of note, specificity was low (23–25%), and sensitivity was the strongest measure (>96%), for all immunohistochemical assays. Interestingly, when 5-year disease-free survival was utilized as the reference standard, the immunohistochemical assays outperformed the ligand-binding assay in sensitivity and accuracy (92–93% vs 73%, and 77–79% vs 68%, respectively). However, the ligand-binding assay showed the strongest specificity at 39% vs 11–15% for the immunohistochemical assays. These results support current immunohistochemical methods for assessing progesterone receptor are superior to the ligand-binding assay.

Our previous publication suggested functional equivalence of the estrogen receptor ready-to-use assays;²⁹ however, the surprising differences in 5-year disease-free survival with the progesterone receptor immunohistochemical assays (Tables 1 and 2; Figure 2) led us to subsequently question if the progesterone receptor results would alter if estrogen

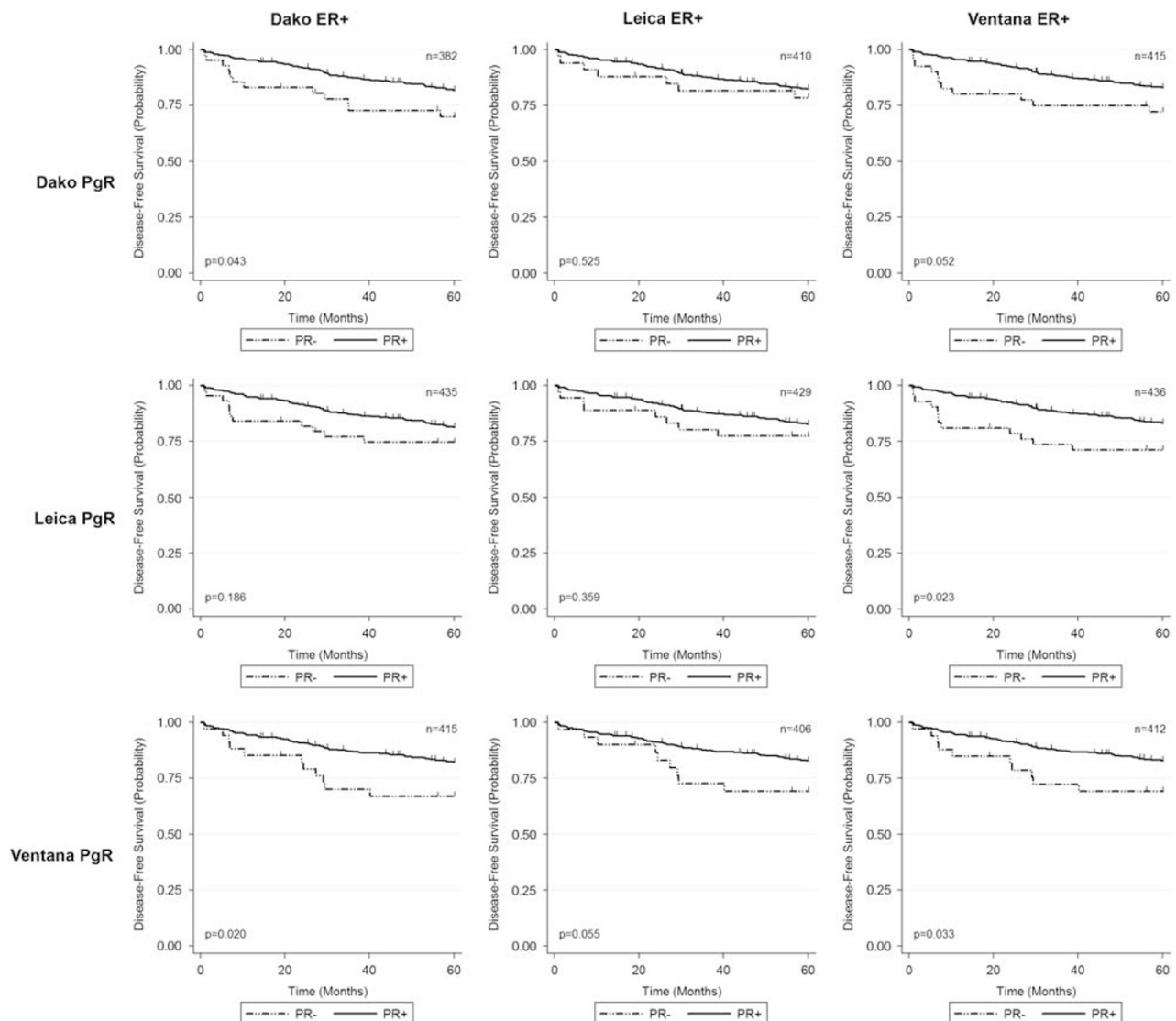


Figure 3 Univariate analysis, including log-rank *P*-values, of progesterone receptor negative (PR-) vs positive (PR+) in estrogen receptor positive cases, as defined by each vendor ready-to-use assay. Total sample size for univariate analysis of each assay is indicated in the upper right corner.

receptor positivity was selected by an alternate vendor assay. To investigate this further, we evaluated the univariate and multivariate survival analyses looking at all combinations of estrogen and progesterone receptor. This exploration led to more surprising results, suggesting that our previous conclusion of functional equivalence with the estrogen receptor ready-to-use immunohistochemical assays may be misleading when it is not utilized in conjunction with progesterone receptor. The Kaplan-Meier curves for all possible combinations are presented in Figure 3, and corresponding univariate proportional hazard estimates are presented in Table 5. Univariate analyses suggests that estrogen receptor measured by the Dako ready-to-use assay yields the strongest prognostic group for the Ventana progesterone receptor ($P=0.023$). Moreover,

the Leica progesterone receptor assay that previously showed no significance now has achieved statistical significance when the Ventana estrogen receptor results were used to define the estrogen receptor-positive group. However, when the estrogen and progesterone receptor combinations were looked at in the multivariate model (Table 6), only the Ventana progesterone receptor was able to maintain significance when adjusting for lymph node status, tumor grade, tumor size and HER2 status. Even more intriguing was that statistical significance was only achieved with the combinations of Leica estrogen receptor positivity with Ventana progesterone receptor ($P=0.026$) and Dako estrogen receptor positivity with Ventana progesterone receptor results ($P=0.037$) (Table 6), suggesting that the estrogen receptor ready-to-use immunohistochemical assays

Table 5 Univariate Cox proportional hazard models for the progesterone receptor Dako, Leica and Ventana immunohistochemical assays with estrogen receptor positivity for each matching vendor assay

	Dako ER+			Leica ER+			Ventana ER+		
	HR	95% Confidence interval	P-value	HR	95% Confidence interval	P-value	HR	95% Confidence interval	P-value
Dako	0.54	(0.29–0.99)	0.046	0.78	(0.36–1.69)	0.526	0.53	(0.28–1.01)	0.055
Leica	0.65	(0.35–1.23)	0.190	0.71	(0.34–1.48)	0.361	0.50	(0.27–0.92)	0.026
Ventana	0.48	(0.25–0.90)	0.023	0.51	(0.25–1.03)	0.059	0.49	(0.25–0.96)	0.036

Table 6 Multivariate Cox proportional hazard models for the progesterone receptor Dako, Leica and Ventana immunohistochemical assays with estrogen receptor positivity for each vendor

	Dako ER+			Leica ER+			Ventana ER+		
	HR	95% Confidence interval	P-value	HR	95% Confidence interval	P-value	HR	95% Confidence interval	P-value
Dako	0.65	(0.31–1.37)	0.261	0.91	(0.36–2.31)	0.845	0.71	(0.33–1.54)	0.390
Leica	0.70	(0.32–1.51)	0.360	0.83	(0.33–2.09)	0.693	0.65	(0.30–1.39)	0.267
Ventana	0.44	(0.20–0.95)	0.037	0.37	(0.16–0.89)	0.026	0.50	(0.22–1.16)	0.090

All models were adjusted for lymph node status, tumor grade, tumor size and HER2 status.

may not be functionally equivalent, as we previously suggested, when looked at in combination with progesterone receptor. As previously mentioned, these findings may partially explain some of the discordance seen with the prognostic/predictive value of progesterone receptor seen in the literature. Further investigation to validate our findings may finally clarify progesterone receptor utility and value in breast cancer management.

In conclusion, despite similar agreement and concordance seen between the vendor-specific progesterone receptor assays, clear differences were noted with regards to 5-year disease-free survival, suggesting that relying on agreement and concordance for the utility of alternate assays may lead to important clinically relevant information being lost. Further investigation into the prognostic, and potentially predictive, value of specific progesterone receptor antibodies is warranted, and should be considered in combination with different estrogen receptor antibody clones.

Acknowledgments

This work was supported by funding provided by the Breast Cancer Society of Canada. Additional support for the construction of the Calgary cohort was provided by Calgary Laboratory Services. In addition, we thank all the vendors—Dako, Leica and Ventana—for providing us with free reagents to complete this study. We would also like to acknowledge Dr Annie Yau and Ms Mie Konno for their assistance in the creation of the database.

Disclosure/conflict of interest

Reimbursement for presentation of results at the Leica Breast Cancer Symposia in London, UK, April 2013 was provided by Leica for ACK. AMM would like to make the following disclosures: consultant and grant recipient (Ventana Medical Systems). The remaining authors declare no conflict of interest. A portion of these results were previously presented at the 2012 San Antonio Breast Cancer Symposium (P1-07-10) as well as at the 2013 Leica Breast Cancer Symposia in London, UK.

References

- EBCTC Group. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005;365:1687–1717.
- EBCTC Group. Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy. *Lancet* 1992;339:1–15.
- EBCTC Group. Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* 1998;351:1451–1467.
- Hammond MEH, Hayes DF, Dowsett M, *et al*. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Clin Oncol* 2010;28:2784–2795.
- Allred CD, Carlson RW, Berry DA, *et al*. NCCN task force report: estrogen receptor and progesterone receptor testing in breast cancer by immunohistochemistry. *J Natl Compr Canc Netw* 2009;7(Suppl 6):S1–S21.

- 6 Banerjee M, George J, Song EY, *et al*. Tree-based model for breast cancer prognostication. *J Clin Oncol* 2004;22:2567–2575.
- 7 Bardou V-J, Arpino G, Elledge RM, *et al*. Progesterone receptor status significantly improves outcome prediction over estrogen receptor status alone for adjuvant endocrine therapy in two large breast cancer databases. *J Clin Oncol* 2003;21:1973–1979.
- 8 Colomer R, Beltran M, Dorcas J, *et al*. It is not time to stop progesterone receptor testing in breast cancer. *J Clin Oncol* 2005;23:3868–3869.
- 9 Dowsett M, Allred DC, Knox J, *et al*. Relationship between quantitative estrogen and progesterone receptor expression and human epidermal growth factor receptor 2 (HER-2) status with recurrence in the arimidex, tamoxifen, alone or in combination trial. *J Clin Oncol* 2008;26:1059–1065.
- 10 Dowsett M, Cuzick J, Wale C, *et al*. Retrospective analysis of time to recurrence in the ATAC trial according to hormone receptor status: an hypothesis-generating study. *J Clin Oncol* 2005;23:7512–7517.
- 11 Fisher ER, Sass R, Fisher B. Pathologic findings from the national surgical adjuvant breast project. Correlations with concordant and discordant estrogen and progesterone receptors. *Cancer* 1987;59:1554–1559.
- 12 Fuqua SA, Cui Y, Lee AV, *et al*. Insights into the role of progesterone receptors in breast cancer. *J Clin Oncol* 2005;23:931–932.
- 13 Gelbfish GA, Davidson AL, Kopel S, *et al*. Relationship of estrogen and progesterone receptors to prognosis in breast cancer. *Ann Surg* 1988;207:75–79.
- 14 Hopp TA, Weiss HL, Hilsenbeck SG, *et al*. Breast cancer patients with progesterone receptor PR-A-rich tumors have poorer disease-free survival rates. *Clin Cancer Res* 2004;10:2751–2760.
- 15 Knutson TP, Lange CA. Tracking progesterone receptor-mediated actions in breast cancer. *Pharmacol Ther* 2014;142:114–125.
- 16 Horwitz KB, McGuire WL, Pearson OH, *et al*. Predicting response to endocrine therapy in human breast cancer: a hypothesis. *Science* 1975;189.
- 17 Osborne CK, Yochmowitz MG, Knight WA, *et al*. The value of estrogen and progesterone receptors in the treatment of breast cancer. *Cancer* 1980;46(12 Suppl):2884–2888.
- 18 Stendahl M, Ryden L, Nordenskjold B, *et al*. High progesterone receptor expression correlates to the effect of adjuvant tamoxifen in premenopausal breast cancer patients. *Clin Cancer Res* 2006;12:4614–4618.
- 19 Thakkar J, Mehta DG. A review of an unfavorable subset of breast cancer: estrogen receptor positive progesterone receptor negative. *Oncologist* 2011;16:276–285.
- 20 Taylor CR. The total test approach to standardization of immunohistochemistry. *Arch Pathol Lab Med* 2000;124:945–951.
- 21 Taylor CR. Focus on biospecimens: the issue is the tissue. Observations on the "NCI-NIST fitness-for-purpose quality assessment and standards development workshop," October 2010. *Appl Immunohistochem Mol Morphol* 2011;19:95–98.
- 22 Sompuram SR, Vani K, Messana E, *et al*. A molecular mechanism of formalin fixation and antigen retrieval. *Am J Clin Pathol* 2004;121:190–199.
- 23 Gown AM. Current issues in ER and HER2 testing by IHC in breast cancer. *Mod Pathol* 2008;21(Suppl 2):S8–S15.
- 24 Yaziji H, Taylor CR, Phil D, *et al*. Consensus recommendations on estrogen receptor testing in breast cancer by immunohistochemistry. *Appl Immunohistochem Mol Morphol* 2008;16:513–520.
- 25 Pekmezci M, Szpaderska A, Osipo C, *et al*. The effect of cold ischemia time and/or formalin fixation on estrogen receptor, progesterone receptor, and human epidermal growth factor receptor-2 results in breast carcinoma. *Patholog Res Int* 2012;2012:947041.
- 26 Tong LC, Nelson N, Tsourigiannis J, *et al*. The effect of prolonged fixation on the immunohistochemical evaluation of estrogen receptor, progesterone receptor, and HER2 expression in invasive breast cancer: a prospective study. *Am J Surg Pathol* 2011;35:545–552.
- 27 Yildiz-Aktas IZ, Dabbs DJ, Bhargava R. The effect of cold ischemic time on the immunohistochemical evaluation of estrogen receptor, progesterone receptor, and HER2 expression in invasive breast carcinoma. *Mod Pathol* 2012;25:1098–1105.
- 28 Fitzgibbons PL, Bradley LA, Fatheree LA, *et al*. Principles of analytic validation of immunohistochemical assays: guidelines from the College of American Pathologists Pathology and Laboratory Quality Center. *Arch Pathol Lab Med* 2014;138:1432–1443.
- 29 Kornaga EN, Klimowicz AC, Guggisberg N, *et al*. A systematic comparison of three commercial estrogen receptor assays in a single clinical outcome breast cancer cohort. *Mod Pathol* 2016;29:799–809.
- 30 Remotti H. Tissue microarrays: construction and use. *Methods Mol Biol* 2013;980:13–28.
- 31 Molnar LA, Dhirani N, Tran K, *et al*. Immunoassay for estrogen receptor does not detect inactivated receptor. *Clin Chem* 1989;35:824–828.
- 32 Wolff AC, Hammond MEH, Schwartz JN, *et al*. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* 2007;25:118–145.
- 33 Allred DC, Harvey JM, Berardo M, *et al*. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Mod Pathol* 1998;11:155–168.
- 34 Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005;37:360–363.
- 35 Akobeng AK. Understanding diagnostic tests 1: sensitivity, specificity and predictive values. *Acta Paediatr* 2007;96:338–341.
- 36 Phillips T, Murray G, Wakamiya K, *et al*. Development of standard estrogen and progesterone receptor immunohistochemical assays for selection of patients for antihormonal therapy. *Appl Immunohistochem Mol Morphol* 2007;15:325–331.
- 37 Dako. Dako ER/PR pharmDx Kit (<http://www.dako.com/ca/download.pdf?objectid=128036002>), 2009.
- 38 Qiu J, Kulkarni S, Chandrasekhar R, *et al*. Effect of delayed formalin fixation on estrogen and progesterone receptors in breast cancer. *Am J Clin Pathol* 2010;134:813–819.
- 39 Arber DA. Effect of prolonged formalin fixation on the immunohistochemical reactivity of breast markers. *Appl Immunohistochem Mol Morphol* 2002;10:183–186.
- 40 Olivetto IA, Truong PT, Speers CH, *et al*. Time to stop progesterone receptor testing in breast cancer management. *J Clin Oncol* 2004;22:1769–1770.