



**POST-SABCS** Issue 3, 2013

## **An International Ki-67 Reproducibility Study**

For more visit [ResearchToPractice.com/5MJCSABCS2013](http://ResearchToPractice.com/5MJCSABCS2013)

**Research  
To Practice®**

## CME INFORMATION

### OVERVIEW OF ACTIVITY

The annual San Antonio Breast Cancer Symposium (SABCS) is unmatched in its significance with regard to the advancement of breast cancer treatment. This unique conference provides many clinicians and scientists from around the world a forum for discussing critical issues in the prevention and management of breast cancer. Therefore, SABCS is targeted by many members of the clinical research community as the optimal time to unveil new clinical data. This creates an environment in which yearly published results from a plethora of ongoing clinical trials lead to the emergence of new therapeutic agents and changes in the indications for existing treatments across all breast cancer subtypes.

Although SABCS currently offers online access to the vast majority of the poster and plenary presentations from the annual meeting, the absence of expert assessment concerning practical applications may yield confusion among community oncologists who are challenged almost daily to appropriately apply a multitude of scientific findings across diverse tumors. Thus, identification of those data sets with immediate or impending relevance to the delivery of quality breast cancer care, in conjunction with professional commentary to address resultant practice ambiguity, may prove vastly beneficial to those physicians who are unable to make the annual pilgrimage to San Antonio. To bridge the gap between research and patient care, this CME activity will deliver a serial review of the most important emerging data sets from the latest SABCS, including expert perspectives on how these new evidence-based concepts can and should be incorporated into on- and off-protocol patient care. This activity will assist medical oncologists, radiation oncologists, breast/general surgeons, nurses and other healthcare professionals in the formulation of optimal clinical management strategies and the timely application of new research findings to best-practice patient care.

### LEARNING OBJECTIVES

- Evaluate the EndoPredict® signature together with a predefined combination of clinicopathologic factors and molecular data as a predictor of late metastases for patients with estrogen receptor-positive and HER2-negative breast cancer.
- Assess the utility of the Oncotype DX® Recurrence Score® in predicting benefit from the addition of paclitaxel to adjuvant doxorubicin/cyclophosphamide for patients with lymph node-negative and estrogen receptor-positive breast cancer treated concurrently with endocrine therapy.
- Compare the performance of the Breast Cancer Index<sup>SM</sup> biomarker to that of the Oncotype DX Recurrence Score and IHC4 score as prognostic factors for distant recurrence of hormone receptor-positive, lymph node-negative primary breast cancer.
- Determine the impact of metabolic syndrome on breast cancer recurrence for patients with high-, intermediate- or low-risk disease as defined by the 21-gene Oncotype DX Recurrence Score assay.
- Appraise the reproducibility of IHC-based Ki-67 biomarker assays and ongoing strategies to increase concordance in analysis and scoring methods.

### ACCREDITATION STATEMENT

Research To Practice is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians.

### CREDIT DESIGNATION STATEMENT

Research To Practice designates this enduring material for a maximum of 1.25 *AMA PRA Category 1 Credits™*. Physicians should claim only the credit commensurate with the extent of their participation in the activity.

### HOW TO USE THIS CME ACTIVITY

This CME activity contains slides and edited commentary. To receive credit, the participant should review the slide presentations, read the commentary, complete the Post-test with a score of 75% or better and fill out the Educational Assessment and Credit Form located at [ResearchToPractice.com/5MJCSABCS2013/3/CME](http://ResearchToPractice.com/5MJCSABCS2013/3/CME).

### CONTENT VALIDATION AND DISCLOSURES

Research To Practice (RTP) is committed to providing its participants with high-quality, unbiased and state-of-the-art education. We assess potential conflicts of interest with faculty, planners and managers of CME activities. Real or apparent conflicts of interest are identified and resolved through a conflict of interest resolution process. In addition, all activity content is reviewed by both a member of the RTP scientific staff and an external, independent physician reviewer for fair balance, scientific objectivity of studies referenced and patient care recommendations.

**FACULTY** — The following faculty (and their spouses/partners) reported real or apparent conflicts of interest, which have been resolved through a conflict of interest resolution process:

Lisa A Carey, MD  
Richardson and Marilyn Jacobs Preyer  
Distinguished Professor for Breast Cancer Research  
Chief, Division of Hematology and Oncology  
Physician-in-Chief  
North Carolina Cancer Hospital  
Associate Director for Clinical Research  
Lineberger Comprehensive Cancer Center  
Chapel Hill, North Carolina

Advisory Committee, Consulting Agreements and Speakers Bureau: Amgen Inc, Bristol-Myers Squibb Company, Genentech BioOncology, Novartis Pharmaceuticals Corporation, Pfizer Inc, Sanofi; Research Support: Genentech BioOncology, GlaxoSmithKline, Sanofi.

Matthew J Ellis, MB, BChir, PhD, FRCP  
Professor of Medicine  
Director, Breast Cancer Program  
Washington University School of Medicine  
St Louis, Missouri

Consulting Agreements: AstraZeneca Pharmaceuticals LP, Genentech BioOncology, Novartis Pharmaceuticals Corporation; Contracted Research: Millennium: The Takeda Oncology Company, Novartis Pharmaceuticals Corporation, Pfizer Inc; Ownership Interest: Bioclassifier LLC, University Genomics Inc.

**EDITOR** — Dr Love is president and CEO of Research To Practice, which receives funds in the form of educational grants to develop CME activities from the following commercial interests: AbbVie Inc, Algeta US, Allos Therapeutics, Amgen Inc, ArQule Inc, Astellas, Aveo Pharmaceuticals, Bayer HealthCare Pharmaceuticals, Biodesix Inc, Biogen Idec, Boehringer Ingelheim Pharmaceuticals Inc, Bristol-Myers Squibb Company, Celgene Corporation, Daiichi Sankyo Inc, Dendreon Corporation, Eisai

Inc, EMD Serono Inc, Foundation Medicine Inc, Genentech BioOncology, Genomic Health Inc, Gilead Sciences Inc, Incyte Corporation, Lilly USA LLC, Medivation Inc, Merck, Millennium: The Takeda Oncology Company, Mundipharma International Limited, Novartis Pharmaceuticals Corporation, Onyx Pharmaceuticals Inc, Prometheus Laboratories Inc, Regeneron Pharmaceuticals, Sanofi, Seattle Genetics, Spectrum Pharmaceuticals Inc and Teva Oncology.

**RESEARCH TO PRACTICE STAFF AND EXTERNAL REVIEWERS —**  
The scientific staff and reviewers for Research To Practice have no real or apparent conflicts of interest to disclose.

This educational activity contains discussion of published and/or investigational uses of agents that are not indicated by the Food and Drug Administration. Research To Practice does not recommend the use of any agent outside of the labeled indications. Please refer to the official prescribing information

for each product for discussion of approved indications, contraindications and warnings. The opinions expressed are those of the presenters and are not to be construed as those of the publisher or grantors.

This activity is supported by educational grants from Eisai Inc, Genentech BioOncology and Genomic Health Inc.

**Hardware/Software Requirements:**

A high-speed Internet connection

A monitor set to 1280 x 1024 pixels or more

Internet Explorer 7 or later, Firefox 3.0 or later, Chrome, Safari 3.0 or later

Adobe Flash Player 10.2 plug-in or later

Adobe Acrobat Reader

(Optional) Sound card and speakers for audio

Last review date: April 2013

Expiration date: April 2014

## Adjuvant chemotherapy for patients with node-positive luminal A breast cancer

To go directly to slides and commentary for this issue, [click here](#).

The December San Antonio Breast Cancer Symposium (SABCS) once again featured a bounty of papers focused on tissue predictors of response to systemic agents, and although none will shake up clinical practice like Dr Soon Paik's legendary 2004 SABCS presentation documenting the predictive value of the 21-gene Recurrence Score® (RS) in tumor samples from patients on the NSABP-B-20 trial of tamoxifen alone or with chemotherapy (CT), on a macro level these translational and clinical findings contribute significantly to our knowledge base and help further divide this disease into specific biologic subsets. However, other than being an outlet for new research data, SABCS is also an exceptional educational event where fascinating sessions often provide new perspectives on patient care.

To that end, the spectacular clinical science symposium on the first day of the conference included a thoughtful and thought-provoking overview by Dr Kathy Albain on the critical and controversial issue of adjuvant treatment for patients with the common luminal A phenotype, defined by Dr Charles Perou and others as having high ER and normal HER2 levels and relatively low proliferation. Dr Albain noted that last year's international meta-analysis demonstrated an overall benefit of CT in ER+ tumors, but the emergence of contemporary assays like the RS has now identified patients with ER+ tumors who are less likely to benefit from CT.

A critical issue in this regard is the patient with a node+ tumor, and to explore this Dr Albain took a unique tack by directly comparing the results of her work (presented in SABCS 2010) evaluating the RS in available tissue from the SWOG-8814 node+ study to the findings from Paik's initial evaluation of the assay in node-negative tumors from B-20. Interestingly, there was a remarkably similar correlation between CT benefit and RS in B-20 and S8814. However, more than 2 years later

### Remarkably Similar Significant Interaction between Chemotherapy Benefit and Recurrence Score in B20 and S8814

Variable/Study	Tamoxifen	Tamoxifen + Chemotherapy
<b>10 yr DRFS B20 N0*</b>		
Low Recurrence Score	97%	93%
High Recurrence Score	60%	73%
<b>10 yr BCSS S8814 N+**</b>		
Low Recurrence Score	93%	88%
High Recurrence Score	54%	73%

\* Paik S, et al. J Clin Oncol 2006; interaction p = 0.038

\*\*Albain KS, et al. Lancet Oncol 2010; interaction p = 0.021

controversy still surrounds the issue. Dr Albain referred to a 2012 [JCO editorial](#) by Dr Dan Hayes in which he enthusiastically supported enrolling eligible patients on the ongoing RxPONDER node+ trial but vehemently objected to withholding CT (and the potential to improve the chance of remaining disease free) in such individuals outside a study setting. As is often the case with education sessions of this type, Dr Albain did not provide a definitive recommendation about the use of tissue predictors in node+ disease, but at [our CME symposium that evening](#) a number of the faculty members, including Drs Hal Burstein, Kim Blackwell, George Sledge and Cliff Hudis, noted that they will selectively obtain a RS in patients with node+, ER+, HER2-negative tumors and a low nodal burden.

After hearing Dr Albain's talk, I invited her to participate in an audio interview, during which she was particularly enthused about the next generation of prospective trials, including TAILORx and RxPONDER (RS in node-negative and node+ settings) and MINDACT and I-SPY 1 (70-gene signature in the adjuvant and neoadjuvant settings), which have the potential to drastically shift how predictive assays are employed in clinical practice. However, until these trials begin to report, oncologists must make these difficult decisions with a less than optimal evidence base and keep abreast of incremental steps forward. In that regard, here's the bottom line on the most recent crop of related SABCS data sets.

### **1. [More on RS in node+ tumors](#)**

A prominent SABCS paper focused on another retrospective/prospective analysis of tissue in patients with ER+ tumors in a large Phase III trial (NSABP-B-28, evaluating AC alone or with paclitaxel in patients with node+ disease) and provided additional evidence that RS can predict outcome in this population. Interestingly, the incremental gain from paclitaxel was not correlated with RS, but the analysis was underpowered to make that determination.

### **2. Molecular profiles to predict risk of delayed recurrence (DR) in patients completing 5 years of adjuvant endocrine treatment**

The SABCS presentation of the ATLAS trial of 5 versus 10 years of tamoxifen was yet another data set demonstrating the critical role of DR in ER+ tumors, and 2 early but encouraging papers reported on novel RT-PCR assays to identify patients at particular risk for these events. The first looked at the ["EndoPredict® Score"](#) in tissue from 2 major Austrian trials, and the second examined the ["Breast Cancer Index"](#) versus RS and IHC4 in 665 primary tumor samples from the TransATAC tissue bank. Although not definitive, it appears that these or other similar assays may one day be able to provide important input on the critical clinical decision of extending endocrine treatment to 10 years or more while also yielding clues about specific genes correlated with the almost mysterious syndrome of DR, particularly in luminal A tumors.

### **3. Another dagger in the heart of Ki-67**

Show this abstract to your friendly local pathologist the next time he or she offers a home brew that can save you the cost of a RS. After reviewing these scary numbers on lack of Ki-67 reproducibility among pathologists, unless perhaps Dr Mitch Dowsett or Dr Matt Ellis is doing your assay, you may want to rethink this approach.

### **4. Correlation between metabolic syndrome (MS) and breast cancer recurrence in luminal A tumors**

This fascinating effort from Dr Albain's group attempted to determine whether the presence of MS is predictive of breast cancer recurrence in RS subtypes. In addition to documenting an overall eye-popping 27% rate of MS (glucose intolerance/diabetes and 2 other factors, including hypertension, dyslipidemia, central obesity and microalbuminemia) among the 332 patients in the study, of great interest was the correlation of MS and recurrence rate in patients with low RS. This intriguing finding suggests that the work of Dr Rowan Chlebowski and many others demonstrating a link between recurrence and metabolic factors like diet, obesity and exercise may be particularly relevant in luminal A tumors.

Next on this series: SABCS papers on CT, including a surprising potentially practice-changing paper on "pseudoadjuvant" treatment for patients with resected local recurrences rendered Stage IV with no evidence of disease.

Neil Love, MD

**Research To Practice**

Miami, Florida

Research To Practice is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians.

Research To Practice designates this enduring material for a maximum of 1.25 *AMA PRA Category 1 Credits™*. Physicians should claim only the credit commensurate with the extent of their participation in the activity.

This activity is supported by educational grants from Eisai Inc, Genentech BioOncology and Genomic Health Inc.

# **An International Ki-67 Reproducibility Study**

## **Presentation discussed in this issue**

Nielsen TO et al. **An international Ki67 reproducibility study.** San Antonio Breast Cancer Symposium 2012;[Abstract S4-6](#).

**Slides from a presentation at SABCS 2012 and transcribed comments from a recent interview with Matthew J Ellis, MB, BChir, PhD, FRCP (2/7/13)**

# **An International Ki67 Reproducibility Study**

**Nielsen TO et al.**

*Proc SABCS 2012;Abstract S4-6.*

Research  
To Practice®

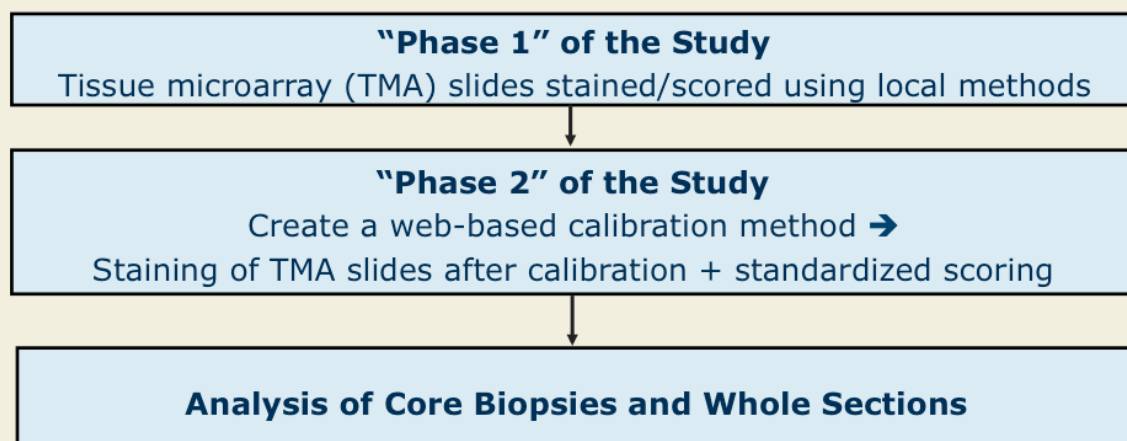
# Background

- Immunohistochemical (IHC) analysis of the cell proliferation marker Ki67 has the potential for utility in breast cancer management for prognostic determinations and in the prediction of treatment response.
- However, the lack of consistent results across laboratories from Ki67 analysis of samples has limited its utility and value as a reliable biomarker.
- In 2010, a working group was assembled to develop and harmonize methodology for Ki67 analysis and scoring and to identify procedures to improve concordance (*JNCI* 2011;103:1656).
- **Study objective:** To conduct an international Ki67 reproducibility study to devise an approach to harmonize Ki67 analytical and scoring methods with the aim of improving concordance.

Nielsen TO et al. *Proc SABCS* 2012;Abstract S4-6.

Research  
To Practice®

# Systematic Study Design



- **Analytical validity:** To determine the extent of reproducibility to which pathologists can reliably quantify Ki67 staining.
- **Endpoints:** Dissemination of analytically valid and meaningful methods for assessing Ki67 and provision of a gold-standard set of web-based calibration cases.

Nielsen TO et al. *Proc SABCS* 2012;Abstract S4-6.

Research  
To Practice®



## Phase 1 Portion of the Study

- Phase 1 determined whether experienced laboratories can deliver consistent Ki67 percentage scores on the same cases using local visual scoring methods.
  - 1-mm TMA cores from 100 breast cancer cases were scored visually by labs using their own scoring methods
- Experiments:
  - Intraobserver (repeat scoring of same TMA slide)
  - Interobserver using a central staining method
  - Interlaboratory observer using local staining methods
- The study included labs from Canada, France, Italy, UK and USA.
  - Study sites included universities, major cancer centers and a national reference lab

Nielsen TO et al. *Proc SABCS 2012*;Abstract S4-6.

Research  
To Practice®

## Phase 1: Results and Lessons Learned

- Intraobserver consistency was good.
  - Six laboratories scored the same 50 cases 3 times.
  - Reproducibility was high, with an intraclass correlation coefficient (ICC) of 0.94 (95% CI: 0.93-0.97).
  - Labs using formal counting methods produced more consistent results than those using visual estimation.
- Interobserver variability using centrally stained TMA slides was problematic:
  - Median Ki67 values across 8 labs ranged from 10% to 28%.
  - Overall ICC was 0.71 (95% CI: 0.47-0.78).
  - At a hypothetical 13.5% cutoff, 32% of cases would be scored as high Ki67 by one lab but low Ki67 by another lab.
- Although intraobserver consistency was good, interobserver variability was problematic.

Nielsen TO et al. *Proc SABCS 2012*;Abstract S4-6.

Research  
To Practice®

## Phase 1: Results and Lessons Learned (Continued)

- Cutoff points are not freely transferable, and local recalibrations against clinical endpoints or reference images are needed.
- Although the staining method added some variability to the results, the major source of Ki67 differences besides patient biology was the scoring method.
  - Estimation versus counting
  - Choice of areas to count
  - Invasive cancer versus other cells
  - Threshold of “brown” considered as “positive” Ki67 staining

Nielsen TO et al. *Proc SABCS 2012*;Abstract S4-6.

Research  
To Practice®

## Phase 2 Portion of the Study

- Phase 2 evaluated whether:
  - Ki67 scorers can be trained in a common visual scoring method that might be transferrable for clinic use.
  - A common reference tool for clinical trial studies can be developed.
- A web-based scoring calibration interface was created from digital images of TMA scores from 9 “training” and 9 “test” cases.
  - TMA slides were centrally stained, representing a range of Ki67 scores.
- A standardized practical scoring method with good internal consistency was chosen.
- Simple instructions with visual examples were developed and provided to 11 laboratories around the world.
- These efforts are continuing with additional labs.

Nielsen TO et al. *Proc SABCS 2012*;Abstract S4-6.

Research  
To Practice®

## Phase 2: Calibration Criteria

- Nine web-based standard images, using a click-tracking application, allowed for the assessment of differences from reference scorers.
- Laboratories with the highest inter- and intralab reproducibility were chosen as reference labs.
  - Their average log2 transformed Ki67 scores = gold standard
- Calibration criteria for study success:
  - Root mean squared differences (RMSE) between volunteer and reference labs among the 9 images: <0.6
  - Maximum absolute difference (MAXDEV) between volunteer and reference lab scores among the 9 images: <1.0

Nielsen TO et al. *Proc SABCS 2012*;Abstract S4-6.

Research  
To Practice®

## Performance Statistics: Training Phase (First Attempt) versus Testing Phase

Score statistic	RMSE (Pass: <0.6)		MAXDEV (Pass: <1.0)	
	Training	Testing	Training	Testing
Mean	0.68	0.41	1.66	0.93
Standard deviation	0.44	0.19	1.26	0.48
Minimum	0.24	0.14	0.35	0.28
Maximum	1.47	0.59	3.82	1.41
Median	0.56	0.49	1.18	1.12

- Differences between training and test cases did not reach statistical significance, possibly due to sample size.
  - For RMSE,  $p = 0.21$
  - For MAXDEV,  $p = 0.22$

Nielsen TO et al. *Proc SABCS 2012*;Abstract S4-6.

Research  
To Practice®

## Phase 2: Study Outcomes and Lessons Learned from Calibration

- Results of calibration testing from 9 participant laboratories and 2 reference labs:
  - 5 of 9 passed testing
- Labs were “trainable,” with improved performance, although improvement was not statistically significant, probably due to sample size.
- Labs differed on the threshold of “brown” staining considered to be positive.
- The following would be added to web-based instructions:
  - Examples of images showing the level of staining that should be considered positive
  - Reminders to consult hematoxylin and eosin (H&E) staining guidelines and not to score ductal carcinoma in situ

Nielsen TO et al. *Proc SABCS 2012*;Abstract S4-6.

Research  
To Practice®

## Future Directions

- If this study is successful:
  - The same scoring system will be applied to core biopsies.
  - Whole sections (with hot-spot issues) will follow later.
  - Clinical utility of analytically valid methods will be confirmed.
  - The levels of expected residual variability in best-practice scoring methods will be defined.
- If this study is unsuccessful:
  - Whether automated platforms and algorithms can deliver consistent results will be tested.
  - Failure would strongly suggest that the Ki67 index by IHC should be used only after internal validation for a given clinical context or as a research tool.

Nielsen TO et al. *Proc SABCS 2012*;Abstract S4-6.

Research  
To Practice®



### **Discussant Comments: An International Study of Ki67 Reproducibility**

This is an elegant study with a strong statistical approach. It points out that pathologists are somewhat "creatures of habit." The overall concordance rate was 71%, which is not particularly strong. With individual research tools, each pathologist becomes somewhat deviant, which is not acceptable in estimating the proliferation rate on a continuous scale. With centrally stained TMA slides, median Ki67 values across labs ranged from 10% to 28%. If we're considering cutoffs in the 15% to 25% range, this will be problematic. Importantly, however, pathologists are trainable.

Improvements were noted from the use of the web-based tool, even though they weren't statistically significant. Bias in choosing tissue sites to count was controlled by the use of TMA slides, but this strategy will introduce some intratumor heterogeneity to the analysis. This will in turn add to the lack of concordance when applied to core biopsies or tissue sections. Also, differences were apparent in identifying the threshold for "brown" staining. Clearly a potential exists for improvement with computer assistance in this field. IHC for Ki67 is currently not reliable for precise measurement of the proliferating fraction in breast cancer. It could be improved with a standardized interpretation method.

***W Fraser Symmans, MD, ASCO 2012***

### **Investigator Commentary: An International Study of Ki67 Reproducibility**

This study illuminates the issues associated with obtaining objective results with Ki67 analysis in terms of the way cells are stained and counted. In order to solve these problems, we're working with a diagnostics company to rapidly develop a rigorous Ki67 test involving the use of a scanner, which will not be so dependent on counting cells with a microscope. With any IHC test variations occur in the quality of the staining, so the staining must be properly controlled for. This can be achieved with the use of an assay kit. If "home brews" by individual labs are used, variation is introduced into the experiments, especially with different incubation periods and the preparation of different antibodies. Thus, different results are generated. Another issue is the heterogeneity of staining as this can be patchy. The same is true for HER2 and ER testing, and the issues mentioned above exist with all IHC tests, not Ki67 only. We believe that the scanner is a good way of solving this problem because it allows for the counting of thousands of cells, which the human eye can't do, and is more objective. If that works out, I believe we will have solved some of these problems. We should be able to push forward with the large validation study that is part of the planned neoadjuvant/adjuvant Phase III ALTERNATE trial, and I am hopeful that in the end Ki67 will enter clinical practice in a more robust and confident manner.

***Interview with Matthew J Ellis, MB, BChir, PhD, February 7, 2013***