Strategies for H-score normalization of preanalytical technical variables with potential utility to immunohistochemical-based biomarker quantitation in therapeutic response diagnostics

William E. Piercealla,∗, Michele Wolfea, Jessica Suschaka, Hua Changa, Yan Chenb, Kam M. Sprotta, Jeffery L. Kutokb,1, Stella Quana, David T. Weavera and Brian E. Warda

aApplied Molecular Diagnostic Group, On-Q-ity, Inc., Waltham, MA, USA
bDepartment of Pathology, Brigham and Women's Hospital, Boston, MA, USA

Abstract. Digital quantitative immunohistochemical analysis of protein biomarker expression offers a broad dynamic range against which clinical outcomes may be measured. Semi-quantitative expression data represented as an H-score is produced by computer generated average intensity of positive staining given weight by the percentage of cells showing positive staining. While patient H-scores vary for biological reasons, variation may also arise from preanalytic technical issues, such as differences in fixation protocols. In this study, we present data on two candidate calibrator nuclear-localized proteins, SNRPA and SnRNP70, with robust and consistent expression levels across breast cancers. Quantitative expression measurement of these two candidate biomarkers may potentially be used to eliminate the effect of differences in preanalytic processing of specimens by normalizing H-scores derived from test biomarkers of interest. To examine the effects of preanalytical fixation variation on biomarker quantitation and potential utility of candidate calibrators to address such issues, 6 surgically-resected human breast cancer patient specimens were divided into 6 portions and fixed under distinct conditions (fixation following resection in formalin for 2 hr, 8 hr or 48 hr, or held overnight at 4 °C in buffered saline prior to formalin fixation for 2 hr, 8 hr, or 48 hr). We find H-score variation between fixation conditions within individual patient’s tumors that were stained for XPF, ATM, BRCA1, pMK2 and PARP1. Most interestingly, detectable expression of SNRPA and SnRNP70 is covariant to test biomarkers under distinct fixation conditions and so these hold the potential for serving as calibration standards for general antigen preservation and reactivity.

Keywords: IHC, theranostics, biomarker, fixation, preanalytical variation

1. Introduction

IHC Biomarker expression levels of FFPE-derived sections have been reported to discriminate therapeutic response in a variety of solid tumor malignancies [1–6]. Creating accurate marker-specific IHC assay cut-off values for prediction of therapeutic response is of great interest, but potential technical issues remain that may hinder correct ascertainment of accurate IHC expression levels of individual biomarkers within patients. One of these potential obstacles may be the effect of variability in time to fixation and formalin fixation time on the protein expression assayed by IHC as scored by digital quantitative image analysis systems [7–11].
While the field has generally employed subjective classification by trained pathologists using standard 1+, 2+, 3+ scoring strategies, advances in digital pathology and image analysis software offer the potential for more objective and quantitative analysis using specific computer scoring algorithms [12–14]. Assays for other molecular analytes utilize standard calibrators to assess overall sample quality and integrity (e.g. RNA employing actin or tubulin measurements). To date, IHC suffers from a lack of such technical standardizing controls. As we move from more subjective assessments of IHC-stained samples to computer assisted quantitative ranking within patient cohorts, corrective criteria for eliminating technical variations need to be applied for establishing models with greatest sensitivity and specificity.

Here we show the variation of patient H-scores derived from tumors that were divided and fixed under different conditions allowing for the possibility of less precise placement within ordered patient population continuums. We have examined two candidate IHC calibrators and show that their expression is covariant with test biomarkers based on the slope of H-scores across various fixation conditions. These criteria are hallmarks of their potential utility as candidate calibrators in normalizing quantitative IHC.

2. Methods

2.1. Patients and tumor samples

Female invasive ductal breast carcinoma tumor samples from 6 individual patients were surgically resected (Pantomics Contract Research Services, Richmond, CA) following guidelines set forth by IRB approval. Patient and individual specimen data is provided in Table 1. Each tumor specimen was divided into 6 parts and each part was fixed in 10% neutral buffered formalin for 2, 8 or 48 hrs, or was held for 24 hours at 4°C in saline before fixation. Specimens were paraffin-embedded and 4 um sections were taken from blocks. All slides were IHC stained within 2 weeks of sectioning.

2.2. Tissue microarrays and immunohistochemistry

Four micron sections from breast cancer microarray TMA 1503 (Pantomics, Richmond, CA) and six individual divided patient samples (36 sections in total) were stained using an automated stainer (Ventana, Tucson, AZ). Sections plus positive control multi-tumor blocks containing representative 0, 1+, 2+ and 3+ breast tissues for each antibody were stained with six biomarkers (ATM, BRCA1, XPF, PARP1, SNRPA and SnRNP70) using optimized protocols. pMK2 was stained manually with a 2 hour primary antibody incubation. Table 2 indicates the specific biomarkers tested, the antibodies utilized, and relevant information with regard to antigen retrieval processes that were employed. Initially, slides were baked at 60°C for 30 minutes in an oven incubator and barcode labels containing specific experimental information were created and placed on each slide. All slides were deparaffinized and following antigen retrieval, incubation was with specified antibodies at optimized dilutions. Detection was via an HRP-conjugated secondary antibody (16 min) and DAB. Slides were counterstained with

| Case | 20101720 | 1803043 | 1004579 | 1003333 | 20102383 | 10037982 |
|---|---|---|---|---|---|---|
| Age | 46 | 60 | 47 | 56 | 57 | 58 |
| Histology | Inv Duct Carc | Inv Duct Carc | Inv Duct Carc | Inv Duct Carc | Inv Duct Carc | Inv Duct Carc |
| Grade | II | II | II | II | II | II |
| T | 2 | 2 | 3 | 2 | 2 | 2 |
| N | 1 | 2 | 1 | 0 | 1 | 1 |
| M | 0 | 0 | 0 | 0 | 0 | 0 |
| Nodal Status | 2/7 nodes+ | 12/18 nodes+ | 3/19 nodes+ | 0/17 nodes+ | 19 nodes+ | 8/31 nodes+ |
| ER | ND | ND | – | + | ND | – |
| PR | ND | ND | – | – | ND | – |
| p53 | ND | ND | +++ | – | ND | + |
| cleb2 | ND | ND | – | – | ND | – |
| Ks-67 | ND | ND | + | – | ND | + |
IHC Staining for Specific Antibodies. Ab-specific staining conditions are listed for each biomarker assessed as part of this study. The optimal conditions for these Abs on the indicated platforms have been previously optimized for signal to noise as well as dynamic expression ranges in study–derived patient specimens as well as control cell-lines.

| Antibody | ATM | BRCA1 | XPF | p53R2 | p53R1 | SNRPA | SNRP70 |
|----------|-----|-------|-----|-------|-------|-------|--------|
| Ab Clone | Epitomics | BioCare | Abcam | Cell Sign | AbD | Abnova | GenWay |
| Conc (µg/ml) | 1:2000 | 1:40 | 1:250 | 1:100 | 1:625 | 1:15000 | 1:125 |
| Final Ab Dil | 60 min, | 60 min, | 60 min, | 120 min, | 60 min, | 60 min, | 60 min, |
| Time/Temp of incubation | 37°C | RT* | 37°C | RT* | 37°C | 37°C | 37°C |

*RT - Room Temperature (21°C+/-0.5°C); **CC1 - Ventana Antigen Retrieval Cell Conditioning Buffer 1 (Tris/Borate/EDTA Buffer pH8.0); ***RiboCC - Ventana Antigen Retrieval Citrate Buffer pH6.0

Staining was conducted on serial sections from patient tumor fragments. In the case of TMAs, H-scores from 2 individual cores per patient were averaged. A core or ROI was considered informative if fifty or greater tumor cells were present for analysis.

2.4. Candidate calibrator identification

A search was instituted to identify candidate calibrators for IHC based staining that would serve to control for technical variation due to tissue processing and stability. Figure 1 illustrates that approx 1300 genes/products were considered [15, 16]. 200 consistently expressed housekeeping genes were nominated for further in silico analyses and attrition of 123 was based on inconsistent protein expression levels across a wide variety of tumor types and normal human tissues (Human Protein Atlas and GeneCards) yielding a candidate list of 77. A second level criteria for selection was based on our principal interest in nuclear-localized DNA damage and repair enzymes, and so focus was narrowed to best candidates with the same localization as biomarkers to which they would serve as calibrators. Dominantly cytoplasmic-localized proteins than a much more time consuming automated macro scoring of the whole section, only a portion of which is applicably tumor. Staining was conducted on serial sections from patient tumor fragments. In the case of TMAs, H-scores from 2 individual cores per patient were averaged. A core or ROI was considered informative if fifty or greater tumor cells were present for analysis.
were eliminated. Also, only proteins with commercially available antibodies were nominated for further study. Collectively, these added elimination criteria resulted in refinement to seven candidate proteins and their expression levels were examined in breast cancer tissue-microarrays (TMAs). Two candidates were identified from this group, ribonucleotide binding proteins SnRNP70 and SNRPA, which had optimal H-score consistency among total patients examined.

3. Results

3.1. Consistent expression levels for candidate calibrators in breast tumors comprising a commercially available TMA

SNRPA and SnRNP70 display narrow ranges of expression across a patient group derived from a commercially available breast cancer TMA (Fig. 2). For comparison purposes, the dynamic range of a test
biomarker of interest is also displayed to illustrate the consistency of H-scores for candidate calibrators across all patients. Note the narrow range of expression for patient specimens comprising the TMA population for SnRNP70 and SNRPA (%CV = 24.7 and %CV = 36.5, respectively) relative to a test biomarker PARP1 (%CV = 69.6) that is expected to yield a more broad dynamic range (% coefficient of variation (%CV) = (Standard deviation (SD)/Mean) X 100.) Fig. 3 shows representative staining from two breast cancer patients utilized in this study. Note IHC staining for expression is robust with ideal signal to noise and staining is specifically restricted to the nucleus for each of the two biomarkers.

3.2. Preanalytical fixation differences and effect on driver biomarker H-scores

In addition to consistency of expression across samples that comprise a study population, a favorable characteristic for an IHC candidate calibrator is covariance with the technical variable for which the candidate calibrator aims to normalize. To that end, we sought to investigate the amount of technical variation that could arise in test biomarkers upon IHC analysis of individual tumors that were divided and fixed under different conditions. Figure 4 indicates staining for a single patient tumor that was divided and fixed under different conditions prior to staining for BRCA1 and phosphoMAPKAP kinase2. The level of fixation has a profound impact on the level of staining (48 hours displays much greater antigen detection sensitivity than 2 hour fixation) and thus when an algorithm is applied to a test specimen set, ordering could be affected by lack of uniformity due to technical handling consideration.

In addition to BRCA1 and phosphoMAPKAP kinase2, XPF, SnRNP70, SNRPA, ATM and PARP-1 nuclear localized IHC signals were assessed in 6 individual patient tumors for 6 individual fixation conditions (immediate immersion for 2 hr, 8 hr, or 48 hr in neutral buffered formalin or held overnight at 4°C.
The H-scores derived from each of the biomarkers generally increased with fixation time. Most interesting here is that the H-score increase with fixation time was consistent for several test biomarkers (ATM, XPF, phosphoMAPKAP kinase2) as well as candidate calibrators SNRPA and SnRNP70 (Fig. 5). Another confirming positive trait for their potential utility as IHC calibrators is that the trendline (slope) for different test biomarkers is indistinguishable between SNRPA and SnRNP70.

While a general trend was observed for all markers relative to fixation conditions, variation was not uniform across all markers (Fig. 5). PARP1 and BRCA1 were shown to be the most variable relative to formalin fixation time, with scores widely ranging from quantitative image analysis. The observation that the trendline (slope) of variation across the different fixation conditions was less similar for BRCA1 and PARP1 may be indicative of a more limited utility for calibrators in removing the technical noise for all test biomarkers with equal efficiency.

A comparison of average H-scores for the patient group \( (n = 6) \) categorized by individual fixation conditions is depicted in Fig. 6. Consistent across different fixations, SNRPA and SnRNP70 displayed more consistent H-score averages than other biomarkers tested as evidenced by a lower %CV. Again, this limited variation relative to other test biomarkers is a hallmark of potential utility as a candidate calibrator analyte. While ATM, phosphoMAPKAP kinase 2, SNRPA, SnRNP70 and XPF showed less dramatic variation among fixation conditions than PARP1 or BRCA1, enhanced antigen detection and H-scores consistently trended higher with longer fixation times.

4. Discussion

The DNA repair nuclear localized proteins examined in this study display substantial variation in computer generated H-scores under varying fixation conditions. These data indicate that uniform fixation...
will eliminate one source of variation for IHC results, and is a necessary precursor to proper patient ranking in studies designed to develop diagnostic algorithms to identify chemotherapy responders and non-responders with high sensitivity and specificity.

In attempting to normalize for pre-analytical technical variation, one would not anticipate that a calibrator that was more or less sensitive to technical variation relative to the biomarker for which it would serve as standard would be the ideal candidate. Rather, the best calibrator would be one whose measured expression moves commensurately by technical variation to the marker for which it normalizes. The data presented here are consistent with both SnRNP70 and SNRPA as being covariant relative to the test markers that have been examined here as measured by trend lines H-scores across different fixation conditions for these breast cancer patient specimens.
Fig. 6. Variation as measured by mean H-score and %CV for given biomarkers and different fixation conditions. The H-scores for each of 6 patients for a given biomarker were averaged per fixation and Std Dev and %CV calculated. Additionally all fixation conditions for all patients were averaged to gain average %CV for a given biomarker.
liminary results also show that R2 values average 3% of patients for each biomarker are affected. Pre-
therapeutic response and resistance models and find that Currently we are testing this in solid tumor chemother-
samples that achieve an acceptable calibrator H-score. Combined with an attenuated ratiometric score for those
could use both a cut-off criteria for inclusion com-
scoring with the calibrator H-score (Test biomarker
driver biomarker H-scores by attenuated ratiometric
formative. Additionally, one could normalize patient
the data set and the specimen removed as unin-
lems a calibrator is technically irrelevant to improving
expected to be expressed at a reasonably constant level
placement within the predictive model is likely to vary
with diagnostic material sources. Calibration of test
patient H-scores for given biomarkers would allow proper
theranostic assessment to be based on biologic variation with minimal technical confounder effects.

The manner in which a candidate calibrator would be employed to improve test data sets upon which ther-
apeutic models of response and resistance remains an open question that demands further testing. One could simply use an H-score cut-off value as a quality met-
ric for patient inclusion. As these marker standards are expected to be expressed at a reasonably constant level across a patient population, samples that do not achieve a reasonable H-score potentially could be considered technically compromised. Thus, for these patient spec-
ims a calibrator is technically irrelevant to improving the

evaluation of estrogen receptor expression is a strong predictive factor of pathologic complete response after anthracycline-based neo-
adjuvant chemotherapy in hormonal-sensitive breast cancer, Breast Cancer Res Treat 124 (2010), 387–391.

While we have chosen to focus on fixation as a tech-
variations with minimal technical confounder effects. Proper theranostic assessment to be based on biologic

References

[1] T. Petit, M. Wilt, M. Velten, J.F. Bodmer, J.P. Feickert, P. Dufour and J.P. Gnaissua. Semi-quantitative evaluation of estrogen receptor expression is a strong predictive factor of pathologic complete response after anthracycline-based neo-
adjuvant chemotherapy in hormonal-sensitive breast cancer. Breast Cancer Res Treat 124 (2010), 387–391.

[2] B.M. Alexander, K. Sprott, D.A. Farrow, X. Wang, A.D. D’Andrea, S.J. Schett, L.C. Collins, D.T. Weaver and J.E. Garber. DNA repair protein biomarkers associated with time to recurrence in triple-negative breast cancer. Clin Cancer Res 16 (2010), 5796–5804.

[3] H.B. Proesch, K. Katreman, O. Modisch, W. Auerbach, M. Dunne, M. Rezaiez and H. Bojar. Predictive biological markers for response of invasive breast cancer to anthracycline/cyclophosphamide-based primary (radio-)
chemotherapy. Anticancer Res 25 (2005), 4615–4621.

[4] S. Gori, A. Silioni, M. Coloeta, J. Ferri, M.G. Mameli, D. Forneschi, L. Stocchi, J. Foglietti, V. Ludovini, E. Minenza, V. De Angelis and L. Cimino, EGFR, pMAPK, pAKT and PTEN status by immunohistochemistry: correlation with clini-
cal outcome in HER2-positive metastatic breast cancer patients treated with trastuzumab. Annals Oncol 20 (2009), 648–654.

[5] K.A. Glanssen, G. Montini and J.C. Swi, EBCCO as a risk stratifier in platinum-based chemotherapy for non-small-cell lung cancer. Curr Opin Pulm Med 13 (2007), 284–289.

[6] S. Siena, A. Sartore-Bianchi, F. Di Nicolantonio, J. Balfour and A. Bandolli, Biomarkers predicting clinical outcome of epidem-
ical growth factor receptor – targeted therapy in metastatic colorectal cancer. Br J Cancer 103 (2010), 1794–1800.

[7] F.L. Nkoy, M.E. Hammond, W. Bees, T. Bischof, B. Rowley, S. Cammali and W. Sausse. Variable Specimen Handling Affects Hormone Receptor Test Results in Women With Breast Cancer: A Large Multihospital Retrospective Study. Arch Pathol Lab Med 134 (2010), 606-612.

[8] M.E. Hammond, D.F. Hayes, M. Dowsett, D.C. Alfred, L.L. Hagerty, S. Blaber, P.L. Fitzgibbon, G. Francis, N.S. Gold-
stein, M. Hayes, D.G. Hicks, S. Lester, R. Love, P.B. Mangu, L. McShane, K. Miller, C.K. Osborne, S. Puick, J. Predoitchin, A. Rhode, H. Sussan, J.N. Schwartz, F.C. Svee, S. Taube, E.E. Toflisov, P. Vallenstein, V. Viale, D. Viscich, T. Wheeler, R.B. Williams, J.L. Wolfliff and A.C. Wolf. American Soci-
ety of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing
of estrogen and progesterone receptors in breast cancer, Arch Pathol Lab Med 134 (2010), 907–922.

[9] T. Oyama, Y. Ishikawa, M. Hayashi, K. Anibe and J. Horiguchi, The effects of fixation, processing and evaluation criteria on immunohistochemical detection of hormone receptors in breast cancer. Breast Cancer 14 (2007), 182–188.

[10] A.C. Wolff, M.E. Hammond, J.N. Schwartz, K.L. Hagerty, D.C. Allred, R.J. Cole, M. Dowsett, P.L. Fitzgibbon, W.M. Hanna, A. Langer, L.M. McShane, S. Paik, M.D. Pegram, E.A. Perez, M.P. Press, A. Rhodes, C. Sluggen, S.E. Taube, R. Tubbs, G.H. Vancer, M. van de Vijver, T.M. Wheeler and D.F. Hayes, American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer, Arch Pathol Lab Med 131 (2007), 18–43.

[11] N.S. Goldstein, M. Ferkowicz, E. Odish, A. Mani and F. Has- tah, Minimum formalin fixation time for consistent estrogen receptor immunohistochemical staining of invasive breast carcinoma, Am J Clin Pathol 120 (2003), 86–92.

[12] M. Dolléde-Filhart, M. Gustavson, R.L. Camp, D.L. Rimoin, J.L. Tonkinson and J. Christiansen, Automated analysis of tissue microarrays, Methods Mol Biol 664 (2010), 151–162.

[13] D. Faratian, C. Kay, T. Robson, F.M. Campbell, M. Grant, D. Rea and J.M. Bartlett, Automated image analysis for high-throughput quantitative detection of ER and PR expression levels in large-scale clinical studies: the TEAM Trial Experience, Histopathology 55 (2009), 587–593.

[14] E. Rexhepaj, D.J. Bromen, P. Holloway, E.W. Kay, A.H. McCann, G. Lundberg, M.J. Duffy, K. Jirstrom and W.M. Gallagher, Novel image analysis approach for quantifying expression of nuclear proteins assessed by immunohistochemistry: application to measurement of estrogen and progesterone receptor levels in breast cancer, Breast Cancer Res 10 (2008), R89.

[15] V. Popovici, D.R. Goldstein, J. Antonov, R. Jaggi, M. Delorme and P. Wirapati, Selecting control genes for RT-QPCR using public microarray data, BMC Bioinformatics 10 (2009), 42.

[16] J.A. Warrington, A. Nair, M. Mahadavappa and M. Try- gansikayas, Comparison of human adult and fetal expression and identification of 355 housekeeping/maintenance genes, Physiol, Genomics 2 (2000), 143–147.

[17] K.A. DeVito, L.A. Charette, D.L. Rimoin and R.L. Camp, Long-term preservation of antigenicity on tissue microarrays, Lab Invest 84 (2004), 1071–1079.

[18] T.W. Jacobs, J.E. Pfreund, I.E. Stillman and S.J. Schmitt, Loss of tumor marker-immunostaining intensity on stored paraffin slices of breast cancer. J Natl Canc Inst 88 (1996), 1054–1059.

[19] J.H. Fergenbein, M. Garcia-Closa, S.M. Hewitt, J. Lis- cowska, L.C. Taktolu and M.E. Sherman, Loss of antigenicity in stored sections of breast cancer on tissue microarrays, Canc Epidemiol Biomark Prevent 13 (2004), 667–672.

[20] C. Blind, A. Kerpennik, M. Pyacza-Gengelbach, G. Fernhnl, N. Deutschmann, M. Dietl, V. Kemn and L. Petersen, Antigenicity testing by immunohistochemistry after tissue oxidation, J Clin Pathol 61 (2008), 79–83.

[21] T.W. Jacobs, A.M. Green, H. Yaws, M.J. Barnes and S.J. Schmitt, HER-2/neu protein expression in breast cancer evaluated by immunohistochemistry. A study of interlaboratory agreement. Amer J Clin Pathol 113 (2000), 251–258.
Submit your manuscripts at http://www.hindawi.com