



Published in final edited form as:

Lab Invest. 2017 March ; 97(3): 329–334. doi:10.1038/labinvest.2016.148.

Proof of the Quantitative Potential of Immunofluorescence by Mass Spectrometry

Maria I. Toki¹, Fabiola Cecchi², Todd Hembrough², Konstantinos N. Syrigos³, and David L. Rimm^{1,4}

¹Department of Pathology, Yale University Medical School, New Haven, CT, USA

²Nantomics Inc., Rockville, Maryland, USA

³3rd Department of Medicine, University of Athens, School of Medicine, Sotiria General Hospital, Athens, Greece

⁴Department of Internal Medicine, Yale University Medical School, New Haven, CT, USA

Abstract

Protein expression in formalin-fixed, paraffin-embedded (FFPE) patient tissue is routinely measured by Immunohistochemistry (IHC). However, IHC has been shown to be subject to variability in sensitivity, specificity and reproducibility, and is generally, at best, considered semi-quantitative. Mass spectrometry (MS) is considered by many to be the criterion standard for protein measurement, offering high sensitivity, specificity and objective molecular quantification. Here, we seek to show that quantitative immunofluorescence (QIF) with standardization can achieve quantitative results comparable to MS. Epidermal Growth Factor Receptor (EGFR) was measured by quantitative immunofluorescence in 15 cell lines with a wide range of EGFR expression, using different primary antibody concentrations, including the optimal signal-to-noise concentration after quantitative titration. QIF target measurement was then compared to the absolute EGFR concentration measured by Liquid Tissue selected reaction monitoring (LT-SRM) mass spectrometry. The best agreement between the two assays was found when the EGFR primary antibody was used at the optimal signal-to-noise concentration, revealing a strong linear regression ($R^2 = 0.88$). This demonstrates that quantitative optimization of titration by calculation of signal-to-noise ratio allows QIF to be standardized to MS and can therefore be used to assess absolute protein concentration in a linear and reproducible manner.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

Corresponding Author: David L. Rimm M.D.-Ph.D., Professor of Pathology, Director, Yale Pathology Tissue Services, Dept. of Pathology, BML 116, Yale University School of Medicine, 310 Cedar St. PO Box 208023, New Haven, CT 06520-8023, Phone: 203-737-4204, FAX: 203-737-5089, david.rimm@yale.edu.

Disclosure/Duality of Interest:

In the last 24 months I have been engaged in the following relationships: I am a consultant to Astra Zeneca, Agendia, Bethyl Labs, Biocept, BMS, Cell Signaling Technology, Cernostics, ClearSight, FivePrime, Genoptix/Novartis, Metamark Genetics, Merck, OptraScan, Perkin Elmer, and Ultivue. I have received honoraria for presentations at Genentech/Roche, Cell Signaling Technology, and Ventana. I hold equity in Metamark Genetics. Cepheid, Genoptix, Gilead Sciences, Pierre Fabre, Perkin Elmer and Nantomics fund research in my lab.

Immunohistochemistry (IHC) analysis of formalin-fixed, paraffin-embedded (FFPE) patient tissue is routinely used to measure common biomarker expression in diagnostic practice. Despite its widespread use, IHC is subject to variability in sensitivity, specificity and reproducibility, and is generally, at best, considered semi-quantitative^{1, 2}. The scoring systems most commonly used, either ordinal or nominal, generate discontinuous, highly subjective data. As such IHC, and protein measurement in general has been considered unreliable as a companion diagnostic test. DNA-based testing, including in situ hybridization for amplification and translocations, and DNA sequencing, has supplanted protein assessment in the companion diagnostic space. Sometimes there is no nucleic acid-based method that can determine the best drug for a given patient and protein measurement would represent a better option, if it were as quantitative and objective as DNA testing.

Toward that goal, we have sought to define methods that can allow protein-based testing to be as quantitative and reproducible as nucleic acid testing. One such method is the AQUA method of automated quantitative immunofluorescence (QIF)³ which calculates protein expression on a continuous scale by quantifying immunofluorescence pixel intensity per unit area. As such, it has been proven as a tool for removing the subjectivity of the traditional scoring system and providing objective and reproducible measurement of targets. However, immunofluorescence (IF) is still considered by some to be only semi-quantitative due to the nature of the detection and measurement methods, which often include an enzymatic amplification of the visualization reagent. When measuring protein, mass spectrometry (MS) is broadly recognized as among the most accurate of methods. It offers high sensitivity, specificity and objective molecular quantification⁴⁻⁷ of proteins in biological samples and, when coupled with selected reaction monitoring (SRM) methods, it can represent a criterion standard for protein measurement. Specifically, the development of mass spectrometry-based Liquid Tissue selected reaction monitoring (LT-SRM) assay, enables the quantification of multiple proteins directly in FFPE patient tumor tissue in a linear fashion across a large range of concentrations^{8, 9}, obviating the need for antibodies or other types of affinity reagents. The LT-SRM assay has been demonstrated to accurately measure expression of Epidermal Growth Factor Receptor (EGFR) in patient tumors¹⁰. Clinical studies are currently defining quantitative cutoffs for EGFR expression that correlate to patient outcomes in various indications. The assay is run in CAP-certified CLIA laboratory and is currently used by physicians for clinical decision making. Here, we determine whether IF can be standardized to mass spectrometry to allow accurate and linear measurement of EGFR as a proof of concept that QIF can be performed as a quantitative method.

Materials and Methods

Cell lines

The cell lines MCF7, HT29, SKBR3, H441, H1355, H1993, H1648, A549, A431, H820, PC9, HCC193, H1975, H2882 and H1650 were purchased from the American Type Culture Collection (Manassas, VA) or donated by other labs. Although the cell lines were not authenticated by our lab, the application herein does not require authentication since the lines are used as standards, not as biological models. Cell lines were selected to represent a range of EGFR expression and different mutation status. H1648, A549, H1355 were

routinely cultured in Dulbecco's modified Eagle medium: Nutrient Mixture F-12 containing 10% fetal bovine serum and 1% penicillin – streptomycin (Life Technologies, Inc., Grand Island, NY). A431 and MCF7 cells were cultured in Dulbecco's modified Eagle medium containing 10% fetal bovine serum and 1% penicillin – streptomycin (Life Technologies, Inc., Grand Island, NY). H820, H1993, H441, H2882, HCC193, PC9, H1650, H1975 cells were cultured in RPMI 1640 medium containing 10% fetal bovine serum and 1% penicillin – streptomycin (Life Technologies, Inc., Grand Island, NY). HT29 and SKBR3 were cultured in McCoys medium containing 10% fetal bovine serum and 1% penicillin – streptomycin (Life Technologies, Inc., Grand Island, NY).

Tissue Microarray

Tissue specimens were prepared in a tissue microarray (TMA) format as previously described¹¹. Representative tumor areas were obtained from formalin-fixed, paraffin-embedded specimens of the primary tumor, and 0.6mm cores from each tumor block were arrayed in a recipient block. A NSCLC array, termed YTMA 332, containing 30 tumor cores from NSCLC patients was constructed from patient cases that had complete surgical resection of primary tumor between 2011 and 2013. Those that had adequate residual tumor from primary site were selected to be cored. YTMA 332 also includes a 2-fold redundancy of 15 cell line cores representing a range of EGFR expression (MCF7, HT29, SKBR3, H441, H1355, H1993, H1648, A549, A431, H820, PC9, HCC193, H1975, H2882, H1650). All tissue was used after approval from the Yale Human Investigation Committee protocol #9505008219, which approved the patient consent forms or in some cases waiver of consent.

Antibody Titration

To optimize the titer for EGFR D38B1 antibody (Suppl. Table 1), a standardization array (YTMA 332) was stained at five concentrations covering two orders of magnitude in serial sections. The antibody was tested at the dilutions 1:50 (0.34 ug/ml), 1:100 (0.17 ug/ml), 1:500 (0.034 ug/ml), 1:1,000 (0.017 ug/ml) and 1:5,000 (0.0034 ug/ml). A titration curve was plotted using the average scores of the highest 10% (representing the signal) and lowest 10% (representing the noise) of the patient tumor cores and the signal-to-ratio. We defined the objectively optimal antibody titer as the one that had the highest dynamic range of signal with the highest signal-to-noise ratio.

Quantitative Immunofluorescence

Fresh 5µm tissue microarray sections were deparaffinized at 60°C for 30 minutes, incubated in xylene (soaking twice for 20 minutes) and rehydrated with ethanol (twice in 100% ethanol for 1 minute, and then in 70% ethanol for 1 minute). Antigen retrieval was performed in a PT module (LabVision, Fremont, CA) with EDTA buffer (Sigma-Aldrich, St Louis, MO) pH 8.0 for 20 minutes at 97°C. After blocking of endogenous peroxidase with 30% hydrogen peroxide in methanol, slides were incubated with a blocking solution containing 0.3% bovine serum albumin in Tris-buffered saline solution and 0.05% Tween solution for 30 minutes at room temperature. Slides were then incubated overnight with a cocktail of EGFR (D38B1, Cell Signaling Technology) rabbit monoclonal antibody and mouse monoclonal cytokeratin antibody clone AE1/AE3 (M3515, Dako). Next, a mixture of Alexa 546 conjugated goat anti-mouse secondary antibody (Molecular Probes, Eugene, OR,

USA) diluted 1:100 in rabbit EnVision reagent (K4003, Dako) was applied to the slides for 60 minutes at room temperature. Cyanine 5 (Cy5) directly conjugated to tyramide (FP1117; Perkin-Elmer) at a 1:50 dilution for 10 minutes was used for target detection and ProLong mounting medium (ProLong Gold; Molecular Probes) with 4,6-diamidino-2-phenyl-indole (DAPI) was used to stain nuclei. Control slides were run for reproducibility alongside each experimental slide staining run. Quantitative measurement of EGFR immunofluorescence was performed using AQUA method (Automated Quantitative Analysis) (Genoptix Medical Laboratory), quantifying fluorescent signal within subcellular compartments, as described previously [3]. A tumor mask was created by binarizing the cytokeratin signal and creating an epithelial compartment. A quantitative immunofluorescence score was calculated by dividing the target pixel intensity by the area of cytokeratin compartment. QIF scores were normalized to the exposure time and bit depth at which the images were captured, allowing scores collected at different exposure times to be comparable. All spots were visually evaluated and cases with staining artifacts or presence of less than 2% tumor area were systematically excluded.

Statistical Analysis

Pearson's correlation coefficient (R) was used to assess the correlation between cell lines' QIF scores and the absolute protein concentration measured by MS-LT-SRM. For the statistical analysis, the average marker score obtained from two available cores of each cell line was used.

Quantitative proteomics assay by mass spectrometry-based LT-SRM

Targeted quantification of EGFR in cell lines by LT-SRM was performed following previously described procedures designed for the quantification of proteomic targets in FFPE tumor tissues [10]. The steps included (i) preparation of liquid tissue lysates from formalin-fixed cell pellets; (ii) addition of known amount of isotope-labeled internal standard peptides into each cell lysate, and (iii) quantitative analysis using LT-SRM. Detailed methods for above steps, for the peptide separation/elution procedures, mass spectrometry parameters, and for the analysis of the results have been previously described¹⁰. The amount of EGFR in each sample (amol per microgram total protein) was calculated from the ratio of analyte to internal standard peak areas, multiplied by the known amount of internal standard spiked and normalized to the total protein amount injected. Samples were analyzed in triplicate.

Results

The EGFR primary antibody used for QIF was initially titrated in a standardization array containing 30 non-small cell lung cancer (NSCLC) patient tumors with a broad range of EGFR expression. To maximize specificity, the primary antibody was tested at five different concentrations covering two orders of magnitude in serial sections, from 0.34 ug/ml (dilution 1:50) to 0.0034 ug/ml (dilution 1:5,000), and the signal-to-noise ratio was calculated by QIF (See Methods). The signal-to-noise ratio was found to be the highest at 0.017 ug/ml primary antibody concentration (Figure 1A).

In parallel, EGFR expression of 15 cell lines with different EGFR expression levels and mutation status was measured by using the same primary antibody concentrations. EGFR measurement by QIF in the cell lines was highly reproducible when the optimal primary antibody concentration was used, $R^2 = 0.86$ (Figure 1B). The two assays were stained, imaged and quantified independently. Representative images of the IF staining on different cell lines are shown in Figure 1C and subcellular localization of the IF signal in A431 cell line is shown in Suppl Fig 1. The QIF scores of the 15 cell lines varied with different EGFR antibody concentration, revealing a wide range of the receptor's expression (Figures 2A–2E). The absolute EGFR concentration in attomols (10–18 mol) per microgram was measured in the same 15 cell line pellets using the LT-SRM assay (Figure 2F). Mass spectrometry measurement confirmed the wide EGFR expression range but also revealed an almost 10-fold EGFR abundance in A431 compared to the other cell lines.

To compare the agreement of the two assays in target measurement, a regression between the IF scores and the absolute protein concentration measured by LT-SRM was plotted (Figure 3A–3E). EGFR expression in A431 exceeded the dynamic range of the IF assay making it the only cell line driving the regression. The high difference of EGFR abundance in this cell line was only shown when the EGFR primary antibody was used at a low concentration, allowing the difference in expression to be distinct (Figure 3E).

When A431 was excluded (Figure 4A–4E), allowing the assays to be compared in the same dynamic range, the best agreement between IF scores and LT-SRM absolute protein concentration was found when the EGFR D38B1 primary antibody was used at the optimal signal-to-noise concentration (0.017 ug/ml), showing a strong linear regression between the two assays ($R^2 = 0.88$) (Figure 4D). The linearity of the agreement decreased when the working concentration moved away from the optimal concentration of the EGFR D38B1 primary antibody. This observation represents an orthogonal method for validation of the concordance between the mass spec and QIF methods.

Discussion

Validation and optimal titration of the reagents used in antibody dependent assays is critical. Subtle differences in antibody concentration have been shown to affect the scoring and the signal-to-noise threshold, and even change biomarker associations with outcomes^{11, 12}. Here, we have demonstrated that quantitative optimization of titration by calculation of signal-to-noise ratio allows IF to be used to assess absolute protein concentration in a linear and reproducible manner with a validated antibody, in spite of the use of an enzymatic amplification in the detection system. While this method requires careful validation with every run and the use of antibody-specific tissue microarrays with cases selected to span the dynamic range of expression, this data supports that IF can be as reproducible and objective as nucleic acid methods.

This study has a number of limitations. Firstly, all the results come from cell line studies and the absolute protein quantification potential of QOF on patient derived tumors has not been validated yet. Secondly, both QIF and MS were performed on FFPE patient tissue and cell lines processed with the standard histopathological methods used in the clinical setting. The

use of IF in other preparations (frozen tissue, etc) would need to be validated separately. Also, this study does not address the issue of pre-analytic variables and loss of epitope availability as function tissue damage prior to stabilization. By depending heavily on cell lines, we control many pre-analytic variables that may affect tissue. There are not yet methods to normalize for pre-analytic tissue damage, in the way “housekeeping” genes are used to normalize some nucleic acid based tests. However, the use of core biopsies followed by immediate fixation represents a method to minimize tissue damage and is also a popular substrate for companion diagnostic tests. Regardless of the method chosen for tissue acquisition, we argue that QIF can accurately measure the epitope present on the slide, with or without the effects of pre-analytic variables.

It is often impossible to determine the cause for the failure of biomarker-driven clinical trials, but the absence of accurate quantitative tests may be one explanation. Currently, protein expression in clinical tissue samples is typically assessed by IHC. IHC is often used in a manner that is subjective with limited reproducibility. Here we show that QIF can be standardized to MS to allow quantitative (accurate, linear and reproducible) measurement of a biomarker protein (EGFR) on glass slides. We hope this work debunks the common perception that IHC or IF can never be quantitative and opens the door to better quality, more accurate, protein-based companion diagnostic testing.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by grants from the NIH including the Yale SPORE in Lung Cancer, P50-CA196530, the Yale Cancer Center Support Grant, P30-CA016359, support from the OKIBEE support grant (KS) and support from Nantomics Inc. and Gilead Sciences Inc. (DR).

References

1. Hsi ED. A practical approach for evaluating new antibodies in the clinical immunohistochemistry laboratory. *Archives of pathology & laboratory medicine*. 2001; 125(2):289–94. [PubMed: 11175655]
2. Allred DC, Harvey JM, Berardo M, et al. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc.* 1998; 11(2):155–68.
3. Camp RL, Chung GG, Rimm DL. Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nature medicine*. 2002; 8(11):1323–7.
4. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*. 2003; 422(6928):198–207. [PubMed: 12634793]
5. Ong SE, Mann M. Mass spectrometry-based proteomics turns quantitative. *Nature chemical biology*. 2005; 1(5):252–62. [PubMed: 16408053]
6. Addona TA, Abbatiello SE, Schilling B, et al. Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. *Nature biotechnology*. 2009; 27(7):633–41.
7. Nilsson T, Mann M, Aebersold R, et al. Mass spectrometry in high-throughput proteomics: ready for the big time. *Nature methods*. 2010; 7(9):681–5. [PubMed: 20805795]

8. Prieto DA, Hood BL, Darfler MM, et al. Liquid Tissue: proteomic profiling of formalin-fixed tissues. *BioTechniques*. 2005; (Suppl):32–5. [PubMed: 16528915]
9. Nishimura T, Nomura M, Tojo H, et al. Proteomic analysis of laser-microdissected paraffin-embedded tissues: (2) MRM assay for stage-related proteins upon non-metastatic lung adenocarcinoma. *Journal of proteomics*. 2010; 73(6):1100–10. [PubMed: 19944198]
10. Hembrough T, Thyparambil S, Liao WL, et al. Selected Reaction Monitoring (SRM) Analysis of Epidermal Growth Factor Receptor (EGFR) in Formalin Fixed Tumor Tissue. *Clinical proteomics*. 2012; 9(1):5. [PubMed: 22554165]
11. McCabe A, Dolled-Filhart M, Camp RL, et al. Automated quantitative analysis (AQUA) of in situ protein expression, antibody concentration, and prognosis. *Journal of the National Cancer Institute*. 2005; 97(24):1808–15. [PubMed: 16368942]
12. Welsh AW, Moeder CB, Kumar S, et al. Standardization of estrogen receptor measurement in breast cancer suggests false-negative results are a function of threshold intensity rather than percentage of positive cells. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2011; 29(22):2978–84. [PubMed: 21709197]

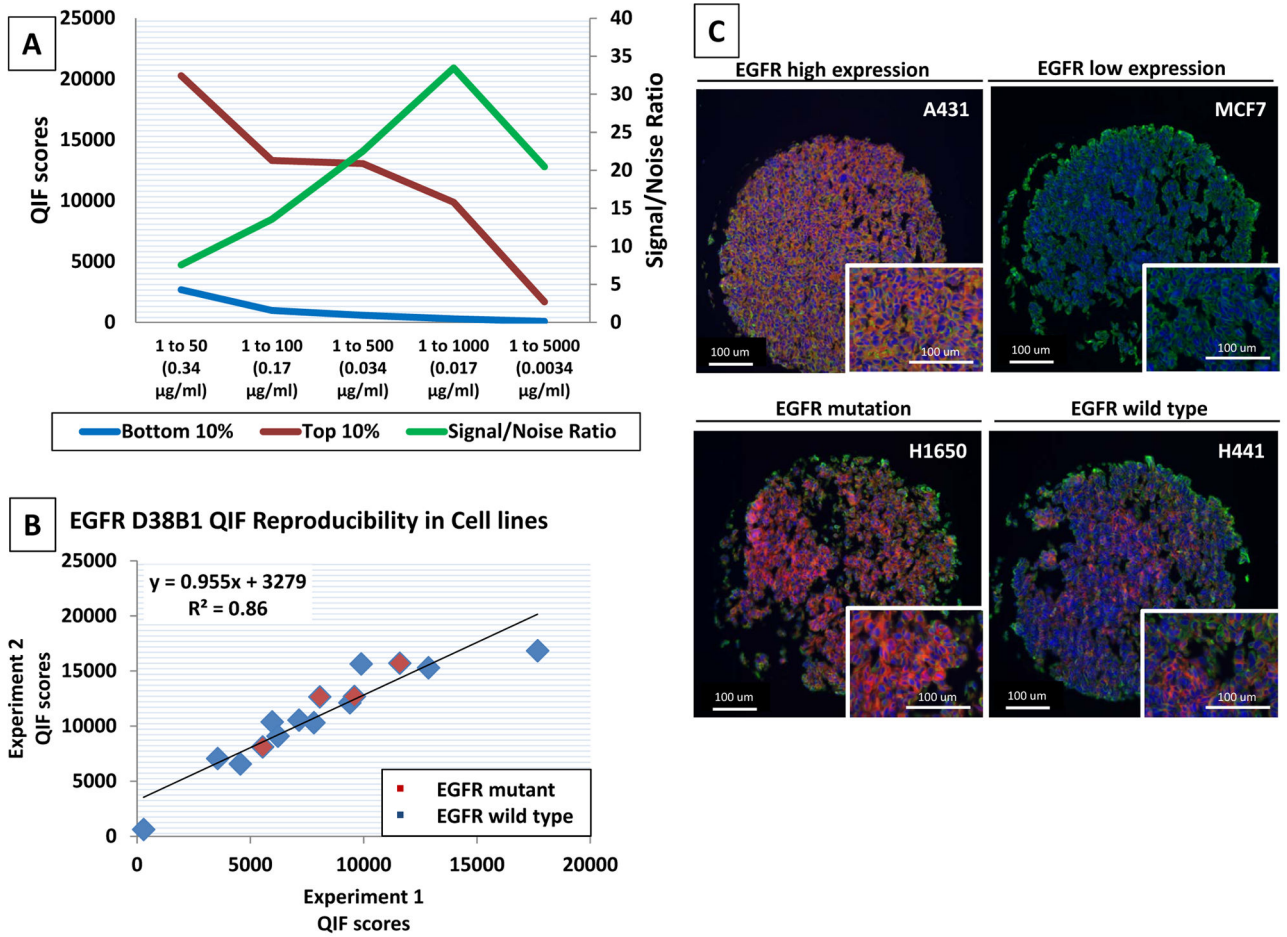


Figure 1. (A) Titration curve of EGFR D38B1 primary antibody plotted at five different concentrations. The blue line shows the average QIF scores of the lowest 10% patient cases included in the standardization array, representing the noise. The red line shows the average QIF scores of the highest 10% patient cases included in the standardization array, representing the signal. The green line is the signal-to-noise ratio for each EGFR D38B1 antibody concentration. (B) Reproducibility of EGFR expression quantified by AQUA in the optimal EGFR D38B1 concentration measured by AQUA in 15 cell lines with different EGFR expression. Red dots represent EGFR mutant cell lines and blue dots represent EGFR wild type cell lines. (C) AQUA images of FFPE cell lines with different EGFR expression level and mutation status. EGFR (red), cytokeratin (green), DAPI (4',6-diamidino-2-phenylindole) (blue). Images are representative of two independent experiments.

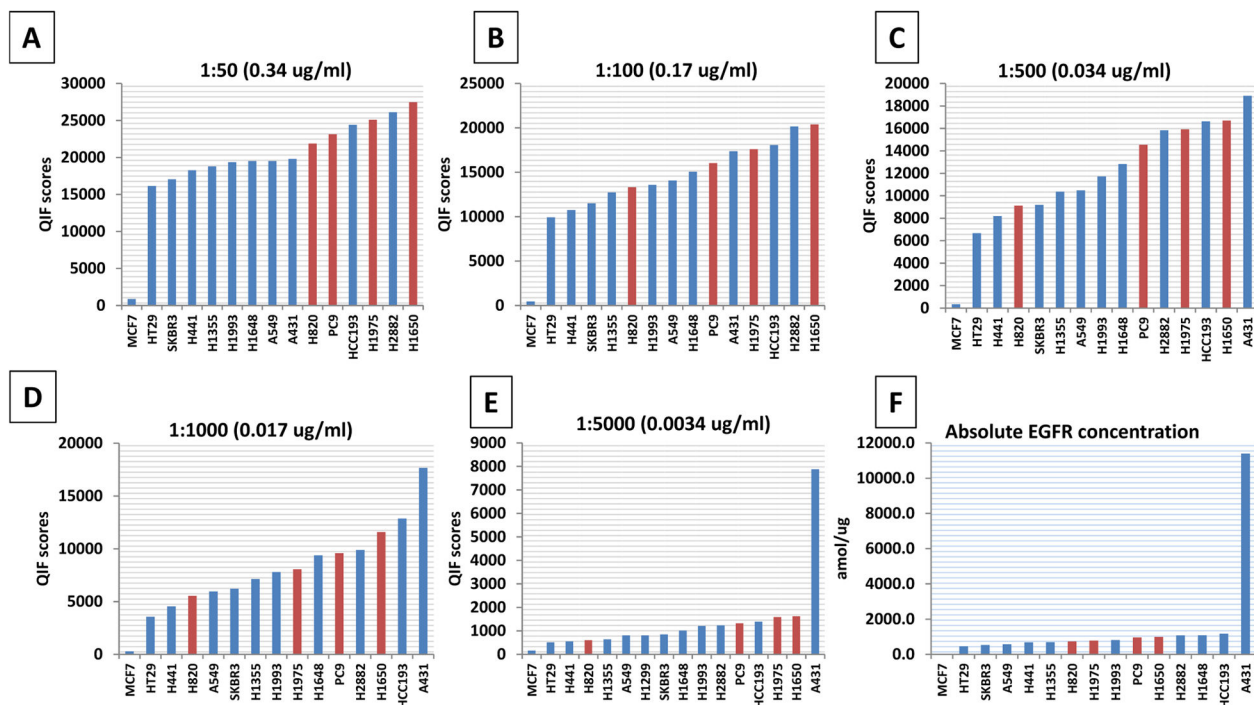


Figure 2. (A–E) QIF scores in the 15 cell lines with different EGFR expression as measured by AQUA[®] at different EGFR D38B1 primary antibody concentrations covering two orders of magnitude in serial sections. (F) Absolute EGFR concentration measured by LT-SRM-MS in 15 cell lines with different EGFR expression. The QIF scoring and the absolute protein concentration by LT-SRM-MS were performed in cores and sections respectively, coming from the same cell pellets. Red bars represent EGFR mutant cell lines and blue bars represent EGFR wild type cell lines.

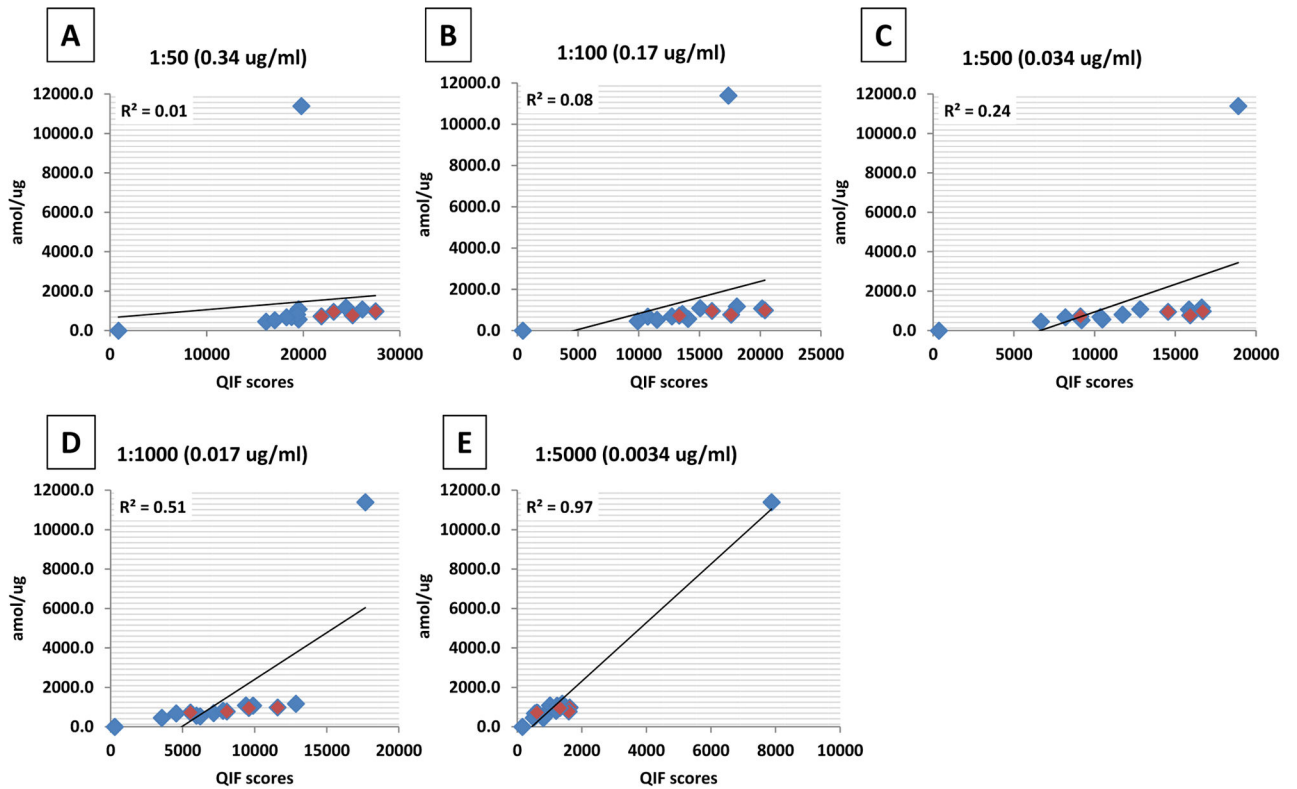


Figure 3.

(A–E) Regression charts between absolute EGFR concentration measured by LT-SRM-MS and the QIF scores at different EGFR D38B1 primary antibody concentrations covering two orders of magnitude in 15 cell lines, with A431 included. The QIF scoring and the absolute protein concentration by LT-SRM-MS were performed in cores and sections respectively, coming from the same cell pellets. Red dots represent EGFR mutant cell lines and blue dots represent EGFR wild type cell lines.

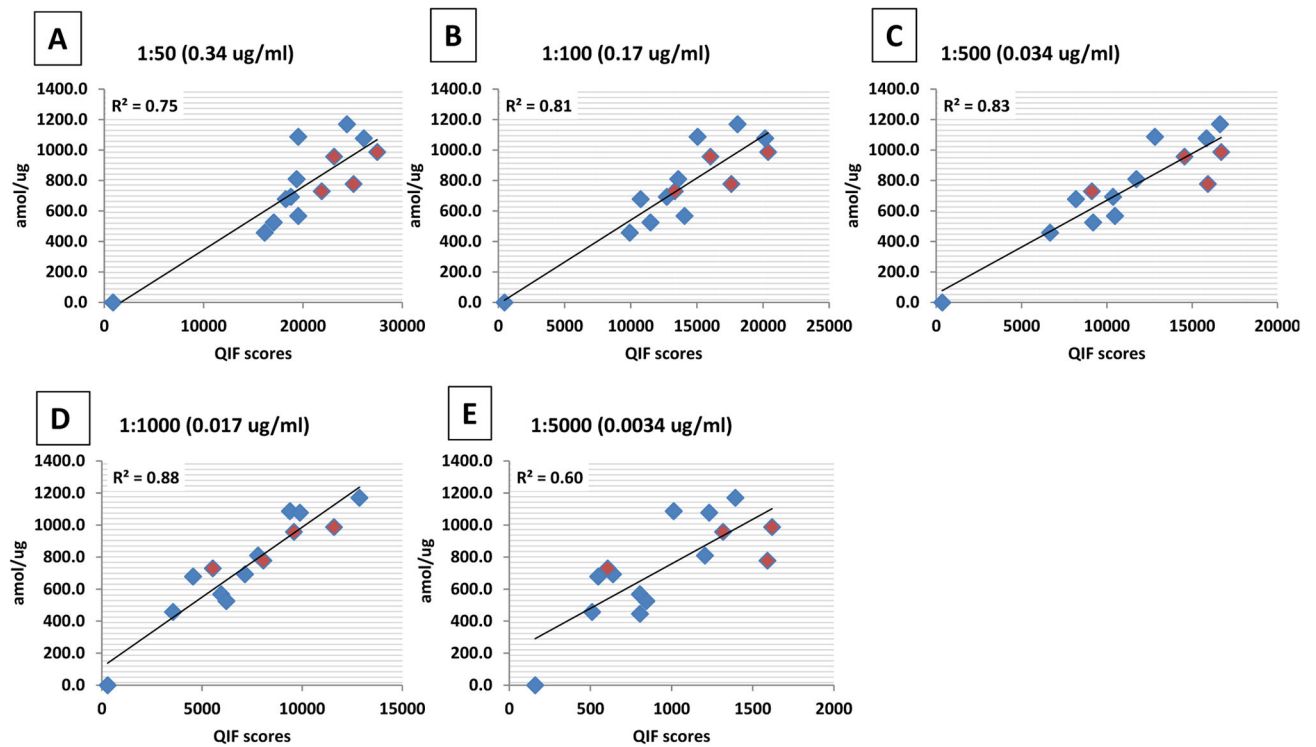


Figure 4.

(A–E) Regression charts between absolute EGFR concentration measured by LT-SRM-MS and the QIF scores at different EGFR D38B1 primary antibody concentrations covering two orders of magnitude in 14 cell lines, with A431 not included. The QIF scoring and the absolute protein concentration by LT-SRM-MS were performed in cores and sections respectively, coming from the same cell pellets. Red dots represent EGFR mutant cell lines and blue dots represent EGFR wild type cell lines.