Validation of a quantitative flow cytometer assay for monitoring HER-2/neu expression level in cell-based cancer immunotherapy products

Britta Randlev, Li-chun Huang, Mitsuko Watatsu, Matthew Marcus, Andy Lin, Shian-Jiun Shih*

Assay Development, Cell Genesys, Inc., 500 Forbes Boulevard, South San Francisco, CA 94404, USA

1. Introduction

Flow cytometry has been an indispensable tool for studying biomarker expression and function in cells and numerous other applications over the past several decades [for review, see [1]]. In recent years, it has also been widely used in the clinical diagnostic and disease progression monitoring [2,3]. Coincident with its increased use, numerous advancements have been made both to the instrument hardware and analysis software of flow cytometry, rendering it more reproducible, accurate, and user-friendly [4]. In addition, the availability of many calibration standards and the software’s ability to meet regulatory compliance requirements have made flow cytometry an appealing addition to Quality Control’s analytical tool box.

GVAX® immunotherapy for prostate cancer is a whole cell therapy comprised of two genetically modified prostate cancer cell lines, CG1940 and CG8711, engineered to secrete granulocyte macrophage-colony-stimulating factor. As part of the matrix of potency assays, CG1940 and CG8711 are tested for the expression level of cell surface HER-2/neu using a quantitative flow cytometer assay. This assay reports the antibody binding capacity of the cells as a measure of HER-2/neu expression using cells immediately after thawing from cryogenic storage. With optimized cell handling and staining procedures and appropriate system suitability controls, the assay was validated as a quantitative assay. The validation results showed that assay accuracy, specificity, precision, linearity, and range were suitable for the intended use of ensuring lot-to-lot consistency of HER-2/neu expression. Assay robustness was demonstrated using design of experiments that evaluated critical assay parameters. Finally, the assay was successfully transferred to a current good manufacturing practice Quality Control laboratory in a separate facility. Since the overall precision of this assay is better than that of ELISA methods and it can be performed with ease and high throughput, quantitative flow cytometer-based assays may be an appropriate immunological assay platform for Quality Control laboratories for characterization and release of cell-based therapies.

138665. Tel: +65 6501 3080; fax: +65 6501 3001.

Abbreviations: QSC, Quantum Simply Cellular; cGMP, current good manufacturing practice; DOE, design of experiment; RSD, relative standard deviation; PE, phycoerythrin.

* Corresponding author at present address: Translational Medicine Research Center, Schering-Plough, 8 Biomedical Grove, #04-01, Neuros Building, Singapore 138665. Tel.: +65 6501 3080; fax: +65 6501 3001.
E-mail address: sj.shih1@gmail.com (S.-J. Shih).

© 2009 The International Association for Biologicals. Published by Elsevier Ltd. All rights reserved.

Keywords:
Quantitative flow cytometry
HER-2/neu
GVAX cancer immunotherapy
To quantify the expression level of surface HER-2/neu per cell, the antibody binding capacity (ABC) of the cells was used. ABC is defined as the number of antibody molecules bound per test object when specific binding sites are saturated [8]. In each assay CG1940 and CG8711 were removed from liquid nitrogen vapor-phase storage and thawed quickly in a warm water bath. Cells were immediately washed to remove cryoprotectants and stained with the anti-HER-2/neu antibody without chemical fixation. The number of HER-2/neu molecules expressed per cell, reported as ABC values, was determined by comparing the fluorescent intensity of bound phycoerythrin (PE)-conjugated anti-HER-2/neu antibody on cells to the fluorescent intensity of the Quantum Simply Cellular (QSC) beads stained with the same PE conjugated anti-HER-2/neu antibody. Every detail of the assay procedure was carefully optimized to achieve the best precision possible while keeping the assay both robust and easy to perform. The assay was initially qualified on the FACS instrument (Becton Dickinson, San Jose, CA) and subsequently further optimized and then validated on the FacsCantoll (Becton Dickenson) equipped with a high-throughput 96-well plate unit to streamline sample handling. Validation of this assay followed regulatory guidance for quantitative assays used to test the potency of therapeutic drugs [9]. The assay performance characteristics, including accuracy, specificity, precision, linearity, range, and robustness showed that this assay is suitable for monitoring HER-2/neu expression levels on current good manufacturing practice (cGMP) lots of CG1940 and CG8711. In addition, the assay was successfully transferred to the cGMP Quality Control laboratory, located at a different site. The data reported here demonstrate that flow cytometer-based assays are suitable for testing therapeutic drugs under GMP guidelines.

2. Materials and methods

2.1. Reagents

Fetal bovine serum (FBS) and all base media were from Hyclone (Logan, UT), L-Glutamine was from JRH (Lenexa, KS), and all antibiotics (penicillin and streptomycin from Invitrogen (San Diego, CA). All reagents for cell staining, solutions for the flow cytometer instrument and the BD™ Cytometer Setup and Tracking Beads were purchased from Becton Dickinson Immunocytometry Systems (BDIS; San Jose, CA.). The PE-conjugated mouse anti-human HER-2/neu was clone Neu 24.7, with a concentration of 5 μg/mL. To evaluate lot-to-lot variation of the PE-conjugated anti-HER-2/neu antibody, multiple lots were analyzed by size exclusion HPLC to confirm nearly a 1-to-1 PE-to-antibody ratio. A qualification procedure was used to demonstrate comparable performance of PE-conjugated antibody to the existing lot in three independent assays prior to using the new lot of antibody. The Spherogem Rainbow Calibration Particles, RCP-30-5A (8 peaks), was from Spherotech (Lake Forest, IL). Based on the long-term stability of these particles established by the manufacturer, they were selected as the calibration system to establish appropriate instrument settings before to using a new instrument for this assay as well as to track daily instrument performance to ensure optical alignment, fluorescence resolution and consistent fluorescence intensity. The geometric mean of each peak was monitored in every assay and was required to pass the predefined limits before QSC or cell staining data could be analyzed. New lots of these calibration particles were required to pass a qualification procedure that empirically demonstrated equivalence between old and new lots in three separate assays before being used.

2.2. Cells

GVAX immunotherapy for prostate cancer (Cell Genesys, South San Francisco, CA) consists of two prostate-cancer cell lines, LNCaP and PC-3, which have been genetically modified with a recombinant replication defective adeno-associated viral vector encoding the human GM-CSF gene. The cells were cultured, harvested and cryopreserved at a concentration of 5.0 × 10⁷ cells/mL, followed by γ irradiation to stop cell proliferation and stored in the vapor phase of liquid nitrogen until use. Jurkat, MCF-7, Hep-3B, OV-90, MRC-5 and KATO III cell lines were obtained from ATCC (Rockville, MD). Jurkat and Hep-3B were maintained in RPMI-1640 with 10% FBS, MCF-7, OV-90, and MRC-5 were cultured in Dulbecco’s modified Eagle’s medium (DMEM) with 10% FBS and KATO III in Iscove’s Modified Dulbecco’s Medium (IMDM) with 20% FBS. All culture media contained 1% L-glutamine and 1% penicillin/streptomycin solution. These cells were cryopreserved at a concentration of 2 × 10⁷ cells/mL prior to use.

2.3. Calibration beads

The qualified QSC beads were obtained from Bangs Laboratories Inc. (Fisher, IN). Each set of QSC beads is consisted of four populations of microbeads, each of which captures a specific amount of mouse IgG molecules due to different levels of goat anti-mouse antibody covalently bound onto the surface. A blank bead population was also utilized. Throughout the validation study, a single lot of QSC beads (Lot 8021) containing the five bead populations with ABC values of 0, 3380, 16286, 77868, and 385810 was used. The ABC values were determined by the manufacturer. Prior to using a new QSC lot in the assay, a new QSC lot would have to be compared to the existing lot and show less than 5% difference in the resulting ABC values.

2.4. Sample preparation for flow cytometry

Vials containing cells were thawed for 2 min in a 37 °C water bath. One milliliter of each cell sample was transferred to another container and slowly diluted with 10 mL phosphate-buffered saline (DBPS) without Ca²⁺ and Mg²⁺ (SAFC Biosciences, Lenexa, KS) containing 1% bovine serum albumin (BSA, from SAFC). Cells were pelleted by centrifugation at 300 × g and the pellets were broken up by flicking the tubes and re-suspended in 30 mL of DBPS with 1% BSA by gently pipetting up and down 3–4 times. Because HER-2/neu expression on the surface of GVAX cells was sensitive to physical manipulation (unpublished results), the number of times a cell pellet was washed and resuspended at each wash step was defined and controlled in the assay procedures. The cells were then adjusted to a final cell concentration of 2.5 × 10⁷ cells/mL, and a 200-μL aliquot was transferred into each of six wells of a 96-well round-bottom plate, three as an unstained control and three for specific antibody staining. To prepare the QSC beads, 25 μL of each bead population were pooled prior to use, and 50 μL of the mixed beads was added to 150 μL of DBPS with 1% BSA and transferred into each of the two wells of the same 96-well plate containing the cells. Cells and beads were pelleted by centrifuging the plate, and the supernatant was aspirated using a 7-mm multichannel plate aspirator manifold (V&P Scientific Inc., San Diego, CA). The PE-conjugated anti-HER-2/neu antibody was diluted to 0.75 μg/mL for the cell staining and 3.5 μg/mL for the QSC bead staining. The concentrations utilized had previously been titrated to ensure binding saturation. Cells and beads were stained with 100 μL of the anti-HER-2/neu working solutions for 30 ± 5 min in the dark at 2–8 °C and washed three times with 200 μL of cold DBPS with 1% BSA. Cells and beads were re-suspended in 200 μL of cold DPBS.
with 1% BSA and left at 2–8 °C in the dark for at least 30 min for antibody equilibration before acquisition by flow cytometry.

2.5. Sample preparation for image analysis

Cells were prepared and stained as per protocols above, except that cells were also counterstained with 0.02 mg/mL Hoechst 33342 nucleic acid stain (Invitrogen, Carlsbad, CA) for 20 min at room temperature. To obtain an optimal density of 10,000 objects per well for image analysis, 100 μL of the cell suspension was transferred to each well of a 96-well flat-bottom plate coated with BD Cell-Tak™ Cell and Tissue Adhesive (BD Bioscience, Bedford, MA). Each population of the QSC beads was stained individually, instead of as a mixture, and 100 μL of each bead population was transferred to each well of the plate. Cells and QSC beads were allowed to adhere to the plate for 30 min prior to image analysis.

2.6. Flow cytometry and ABC value determination

Geometric area and channel data from cells and beads were collected using FACSCanto II™ with High Throughput Sampler HTS station (BDI). The instrument was configured for PE excitation at 488 nm and the detection PMT 556LP dichroic mirror with a 585/42 band-pass filter. The results of acquired samples were displayed and analyzed using BD FACSDiva™ Software, version 6.0. Ten thousand events per well was acquired. A linear regression curve was constructed by plotting the acquired fluorescence intensities from the QSC beads, expressed in geo mean area, against the back-calculated ABC values using an Excel spreadsheet—QuickCal—provided by Bangs Laboratories for each lot of the QSC beads. The ABC value of each test article was then determined by comparing its fluorescent intensity against the standard curve. Every calculation formula in the QuickCal spreadsheet was verified for mathematical accuracy and all spreadsheet cells containing formulas were locked to prevent changes prior to using it for calculating the ABC values in assay validation and by the cGMP Quality Control laboratory.

2.7. High content image analysis

High Content Screening (HCS, also known as High Content Analysis) was performed using the Celldomics ArrayScan 4.5 high-content screening system (Thermo Fisher Scientific, Pittsburgh, PA) which combined automated fluorescence microscopy with multiparameter quantitative image analysis. The Target Activation Assay Algorithm was used for image acquisition. A total of 500 objects per well were imaged using 20 ×/0.4NA objective. Nuclear staining by Hoechst dye was detected in Channel one and used to identify valid objects and define the surface area of each cell. Surface expression levels of HER-2/neu were quantified with Channel two (with PE light path 488/em > FITC dichroic > 570LP/em). The exposure time was 4.611 ms for Channel one and 205.8 ms for Channel two. A Spot Detector Assay Algorithm was used for bead image acquisition. A total of 500 objects per well were imaged using the 10 ×/0.3NA objective. Only Channel two was used for the QSC beads, with the same exposure time as for cells. Fluorescence intensities of cells and beads were reported in pixels.

2.8. Validation study design

2.8.1. Accuracy

Due to lack of a cellular reference with known HER-2/neu expression, the assay accuracy of the flow cytometer assay was inferred by a combination of orthogonal methods. First, taking advantage of the QSC beads with certified ABC values, the recovery of the ABC value for each bead population was used to indicate the assay accuracy. Second, by utilizing several cell lines with known relative HER-2/neu expression, the correlation between the measured ABC values and the expression level estimated by an orthogonal method, the Celldomics High-Resolution Cellular Imaging system, was evaluated. The cell lines evaluated included MCF-7, Hep-3B, OV-90, MRC-5 and KATO III cells, in addition to CG1940 and CG8711, since the tested cell lines are known to have HER-2/neu expression levels that are likely to be within the calibration range of the QSC beads. A good correlation between the readouts from the two platforms for both QSC and cell samples would imply a high level of accuracy; therefore, the acceptance criterion for coefficient of correlation (r) was set as ≥0.80 for QSC beads and ≥0.70 for cells. These acceptance criteria were selected as the minimally acceptable indication for the assay's ability to distinguish product lots with HER-2/neu expression above or below the specification. A lower value was assigned to cells since the cell lines had only high or low HER-2/neu expression levels, unlike QSC beads, which have ABC values evenly spaced across the detectable range and because the cell samples tended to have higher variation than QSC beads. Third, CG1940 and CG8711 cells were stained with serially diluted PE-conjugated anti-HER-2/neu antibody to generate cells with a reduced amount of bound antibody, mimicking cells with less HER-2/neu expression. The relationship between the measured ABC values and the staining antibody titers that were below the saturation concentration further inferred assay accuracy.

2.8.2. Specificity

The assay specificity was demonstrated by the ability of the HER-2/neu antibody to distinguish different levels of HER-2/neu expressed on various cell lines, including Jurkat cells, a T-cell leukemia cell line (ATCC, clone E6-1) that does not express HER-2/neu (unpublished results).

2.8.3. Precision

The assay precision was evaluated at three levels. The repeatability precision was determined in two aspects: one for the triplicate wells of the same sample preparation and another for the duplicate vials from the same lot of final product. The intermediate precision was calculated based on the ABC results from six experiments performed by two analysts over three assays, with each assay testing two lots each of CG8711 and CG1940. Reproducibility was evaluated by comparing the intermediate precision of one laboratory to that of another laboratory using the same lots of cells. The acceptance criterion for precision was that the relative standard deviation (RSD) for all results must be lower than 10%. This precision cutoff value was selected based on the assay capability required to detect and reject CG1940 and CG8711 lots with HER-2/neu expression levels below specification.

2.8.4. Linearity and range

The assay linearity and range were assessed using the QSC beads by comparing the back-calculated ABC value of each bead population to their respective ABC values supplied by the vendor in six independent assays. To meet acceptance criteria, the recoveries of the back-calculated ABC values for the beads were required to be within 75–120%, based on assay development data, and the coefficient of determination (R²) had to be greater than or equal to 0.90 (this value was chosen because of the log transformation of the data involved in back calculation).

2.8.5. Robustness

The first step in the evaluation of assay robustness was to identify potential critical assay parameters and group them into different categories based on initial assay development studies,
their potential interaction, and the feasibility of performing the designed experiment with a reasonable sample load that was similar to actual assay conditions. A high and a low value for each parameter to cover the operational range specified in the assay protocol were defined. For a highly critical factor such as antibody staining time, a linear range with multiple time points were used to better assess the impact of the parameter on assay readout. The range of each evaluated parameter is described in Section 3. The Custom Design of Experiment function of the JMP software (SAS, San Jose, CA) was used to design the robustness experiments and its multi-variance analysis function (Fit Model) was used to analyze data. The criterion for robustness was that the tested parameter should have either no statistical significance ($p$ value $>0.05$) or no practical significant impact on the final results. Lack of practical significance could be established when the RSD between the results from the high and low ends of the tested range was smaller than the assay variation defined by RSD from the precision study.

2.9. Statistical analysis of data

All statistical analyses were performed using JMP software (SAS, San Jose, CA), version 7.0.

3. Results

3.1. Accuracy

Since an established cell reference standard with a known amount of HER-2/neu molecules has not been identified, multiple orthogonal approaches were evaluated to establish assay accuracy. The initial approach was to verify the different $ABC$ values on the QSC standard beads by the Cellomics’s high content image analysis system. This method is orthogonal to the flow cytometry method and is widely used for quantitative purposes for subcellular events [10–13]. The readouts from the Cellomics’s pixel values correlated well with $ABC$ values determined by the flow assay, with a coefficient of correlation of $r = 0.98$ (data not shown). Next, the correlation between the two methods was evaluated for cell samples using two approaches. First, staining with serially diluted antibodies was utilized to mimic differential antigen expression. The differentially stained cells were then assessed by Cellomics and FACSCanto II systems. The coefficients of correlation for the two methods were 0.99 for CG1940 and 1.0 for CG8711 (Fig. 1A and B). Second, based on reported qualitative differential HER-2/neu expression levels, seven different cell lines (Table 1) were selected for evaluation using the two assay systems. The coefficient of correlation between the two systems for these seven cell lines was 0.86 (Fig. 1C). Both tests passed the acceptance criterion for cell staining (coefficient of correlation, $r$, must be larger than or equal to 0.70). Given the correlation between these two orthogonal methods and the assay precision, linearity and specificity (presented below), the accuracy of this quantitative flow cytometer assay in measuring cell surface marker expression can be inferred.

3.2. Repeatability precision

Repeatability was evaluated for well-to-well (replicate) repeatability and vial-to-vial (sample preparation) repeatability. The well-to-well repeatability was determined from the readings from three replicate wells of the same sample preparation. The vial-to-vial repeatability was determined from the readings of two vials representing independently processed (i.e., thawed, diluted, stained, washed) samples of the same lot. The RSD for well-to-well and vial-to-vial repeatability ranged from 0.2% to 2.1% for CG1940, and from 0.2% to 3.6% for CG8711. The RSD of the duplicate sample results from each assay ranged from 0.2% to 2.7% for CG1940 HER-2/neu, and from 0.0% to 4.6% for CG8711 HER-2/neu (Fig. 2). All data met the predefined acceptance criterion for the assay validation (RSD must be lower than 10%).

![Fig. 1](image-url) Correlation between the Cellomics imaging method and ABC determination by flow cytometer. (A) CG8711 stained with decreasing amounts of anti-HER2/neu antibody. Correlation factor = 1.00, (B) CG1940 stained with decreasing amounts of anti-HER2/neu antibody. Correlation factor = 0.99. (C) Various cell lines with different HER2/neu expression: △: CG1940; •: CG8711; □: MCF-5; ○: KATO III; ●: OV-90; ■: MCF-7; ●: Hep 3B. Correlation factor = 0.86. The solid line in each panel represents the linear regression line and the two flanking dotted lines indicate the 95% confidence interval of the regression line.
3.3. Intermediate precision

The intermediate precision was evaluated at three levels: overall intermediate precision, day-to-day precision and analyst-to-analyst precision. The overall intermediate precision was evaluated using the RSD from the combined results from three independent assays conducted by each of the two analysts \((N = 6)\). The day-to-day variability was determined by the RSD over 3 days \((N = 3)\) for each analyst. The analyst-to-analyst variability was evaluated by RSD between the mean results of the two analysts.

### Table 1

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Description</th>
<th>HER-2/neu level</th>
</tr>
</thead>
<tbody>
<tr>
<td>CG1940 ([27,28])</td>
<td>Prostate carcinoma</td>
<td>Low</td>
</tr>
<tr>
<td>CG8711 ([28])</td>
<td>Prostate carcinoma</td>
<td>High</td>
</tr>
<tr>
<td>Hep 3B ([29])</td>
<td>Hepatocellular carcinoma</td>
<td>Mid</td>
</tr>
<tr>
<td>KATO III ([30])</td>
<td>Gastric carcinoma</td>
<td>High</td>
</tr>
<tr>
<td>MCF-7 ([31,32])</td>
<td>Breast carcinoma</td>
<td>High</td>
</tr>
<tr>
<td>MRC-5 ([33])</td>
<td>Normal fibroblast</td>
<td>Mid</td>
</tr>
<tr>
<td>OV-90 ([34,35])</td>
<td>Ovarian carcinoma</td>
<td>Low</td>
</tr>
</tbody>
</table>

### Fig. 2

Distribution of RSD for repeatability (A) RSD values for well-to-well repeatability (calculated from triplicate wells) and (B) RSD values for vial-to-vial repeatability (calculated from duplicate vials). The study was conducted by two analysts over the course of 3 days. Every assay used two lots each of CG1940 and CG8711 and two QSC preparations (prepared by independently mixing five bead populations of the same lot) to measure the effect of variables such as production lots of the cell sample and QSC preparations on repeatability.
An acceptable intermediate precision has been demonstrated for the assay (Table 2A). The RSD for overall intermediate precision (variation across days and analysts) was less than 7% for CG1940 and less than 4% for CG8711. The RSD for day-to-day variation after averaging the results from two analysts was less than 4% for CG1940, and less than 3% for CG8711. The analyst-to-analyst variation was initially assessed by ANOVA to detect analyst specific bias. A statistically significant difference was found in most, but not all, cases between the analysts. For CG1940, the results of Analyst 2 were higher than those of Analyst 1, while for CG8711, the results of Analyst 2 were lower than those of Analyst 1. The extent of the analyst-to-analyst variation was evaluated by RSD between the mean results of Analyst 1 and 2. The largest RSD was 7.0% for CG1940 and 2.7% for CG8711 (Table 2B). All results met the acceptance criteria of RSD <10%.

### 3.4. Reproducibility

Inter-laboratory precision was studied after the assay transfer from the Assay Development (AD) laboratory in South San Francisco, CA to Quality Control (QC) laboratory in Hayward, CA. The assay was performed by two QC analysts over the course of 3 days in the QC laboratory. The ABC values of the same CG1940 and CG8711 lots obtained from the QC and AD laboratories were compared (Fig. 3). There was no statistically significant difference between the laboratories. In addition, the largest RSD for the well-to-well repeatability precision was 2.3% for CG1940, 4.2% for CG8711; the largest RSD for the vial-to-vial repeatability precision was 2.1% for CG1940 and 4.5% for CG8711 in QC results. Both are comparable to the results obtained in the AD laboratory. Between the analysts, the largest RSD for the vial-to-vial repeatability precision was 1.1% for CG1940 and 0.6% for CG8711. Therefore, the QC laboratory demonstrated comparable intermediate precision to that of the AD laboratory. Taken together, good reproducibility was demonstrated for the assay.

### 3.5. Linearity

Linearity was evaluated as the recovery of ABC value for each individual QSC bead population, which served as a surrogate marker, and no ABC value was detected for HER-2/neu on Jurkat cells, a T-cell leukemia cell line, were employed as the negative cell line for the marker, and no ABC value was detected for HER-2/neu on Jurkat cells.

#### Table 2A

<table>
<thead>
<tr>
<th>Lot</th>
<th>QSC Prep</th>
<th>Analyst 1</th>
<th>Analyst 2</th>
<th>Mean (3 days)</th>
<th>SD (3 days)</th>
<th>%RSD (3 days)</th>
<th>Overall mean (N = 6)</th>
<th>Overall SD (N = 6)</th>
<th>Overall %RSD (N = 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CG1940</td>
<td></td>
<td></td>
<td></td>
<td>Mean (3 days)</td>
<td></td>
<td></td>
<td>Overall mean (N = 6)</td>
<td>Overall SD (N = 6)</td>
<td>Overall %RSD (N = 6)</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>Mean (3 days)</td>
<td></td>
<td></td>
<td>Overall mean (N = 6)</td>
<td>Overall SD (N = 6)</td>
<td>Overall %RSD (N = 6)</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>Mean (3 days)</td>
<td></td>
<td></td>
<td>Overall mean (N = 6)</td>
<td>Overall SD (N = 6)</td>
<td>Overall %RSD (N = 6)</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>Mean (3 days)</td>
<td></td>
<td></td>
<td>Overall mean (N = 6)</td>
<td>Overall SD (N = 6)</td>
<td>Overall %RSD (N = 6)</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>Mean (3 days)</td>
<td></td>
<td></td>
<td>Overall mean (N = 6)</td>
<td>Overall SD (N = 6)</td>
<td>Overall %RSD (N = 6)</td>
</tr>
</tbody>
</table>

The mean values are in bold.

#### Table 2B

Intermediate precision, Analyst-to-analyst variation for CG1940 and CG8711 results.

<table>
<thead>
<tr>
<th>Lot</th>
<th>QSC Prep</th>
<th>Analyst 1</th>
<th>Analyst 2</th>
<th>Mean</th>
<th>SD</th>
<th>RSD (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CG1940</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>Mean (3 days)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>Mean (3 days)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>Mean (3 days)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>Mean (3 days)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The mean values are in bold.

#### sample. The recoveries and precision of the back calculated values for each bead population as well as the linear regression between the expected and observed values across the range defined by the QSC beads were studied against the pre-set criteria. The results demonstrated acceptable levels of linearity in assays performed by two analysts across three days. The recoveries of the back-calculated ABC values for the beads ranged from 81.7% to 120.0%, meeting the pre-set acceptance criterion of 75%–120%. The expected values plotted against the observed value (Fig. 4) gave an R² value of 1.00 for both QSC preparations.

#### 3.6. Range

The range of the assay was defined by the ABC values from the lowest to the highest bead population. As shown in Section 3.5, acceptable levels of recovery and precision (with the largest RSD for repeatability being 3.4%) were demonstrated for the ABC range of 3880–385,810 for the QSC bead, lot 8021, which encompassing the expected ABC values for CG8711 (80,000–130,000) and CG1940 (40,000–70,000).

#### 3.7. Specificity

Specificity was verified by staining cell lines known to have different levels of HER-2/neu expression. Jurkat cells, a T-cell leukemia cell line, were employed as the negative cell line for the marker, and no ABC value was detected for HER-2/neu on Jurkat cells.
Fig. 3. Reproducibility of HER-2/neu Expression. (A) Results of ABC value for CG1940 sample. (B) Results for CG8711 sample.

(Her-2/neu) HER-2/neu’s presence on both CG1940 and CG8711 was confirmed by ABC values (>40,000), well above unstained CG1940 and CG8711 cells (<10,000, Fig. 5). These results demonstrate specific binding of anti-HER-2/neu antibody to HER-2/neu on CG1940 and CG8711.

3.8. Robustness

Factors evaluated for robustness (Table 3) were studied in six sets of experiments, including a side-by-side comparison of QSC preparations, an antibody incubation time course and four sets of experiments designed by Design of Experiment (DOE) using JMP software.

3.8.1. QSC preparations

Because QSC beads were mixed immediately before use in each assay, variations between bead preparations were examined. The experiment was conducted by two analysts over 3 days. First, Matched Pair analysis (JMP 6.0) was conducted to compare the ABC values of the same sample calculated using two QSC standard curves. There was a statistical significant difference for the values of CG1940 HER-2/neu ($p = 0.0010$) and CG8711 HER-2/neu ($p < 0.0001$) (Fig. 6A). However, the actual difference in ABC readout between the preparations was small and the largest RSD were 1.3% for both cell lines, thus such a difference was not considered as practically significant.

3.8.2. Antibody incubation time

The effect of antibody incubation time on the ABC determination was studied in a time course from 25 to 35 min. No statistically significant trend was seen for CG8711 ($p = 0.2668$); however, a significant reduction in ABC values at the longer incubation time was confirmed for CG1940 (Fig. 6B), where the RSD between the
mean ABC values at the 25 and 35 minute incubation time in this experiment was 6.5%, smaller than the largest intermediate precision of the assay (7.0%). Therefore, although the incubation time was a statistically significant factor for the ABC determination for CG1940, the range of 25–35 min was considered robust for the assay.

3.8.3. Stability of the cell lines after thaw and after wash

The post-thaw and -wash stability study examined the effect of sample storage time, 0–3 h on ice before antibody staining, on the surface ABC expression. Two conditions were independently examined: (1) delayed processing after thaw but before wash, and (2) delayed processing after wash but before staining.

The study showed a product-dependent loss of expression (Fig. 7). CG8711 appeared to be sensitive to the storage time with a bigger loss of HER-2/neu expression with longer storage both after thaw (13%, \(p < 0.0001\)) as well as after wash (6%, \(p < 0.0001\)). In contrast, CG1940 showed a minimal loss of HER-2/neu expression over the tested holding time. To minimize variability in determining the expression level of HER-2/neu, the hold time post thaw and post wash was specified to be within 1 h.

3.8.4. DOE 1 for cell density, mixing and dispersion, and hold time

The effect of cell density during staining, the number of mixing steps and dispersion of cell suspension during cell staining and wash, and holding time after completion of staining the cells with antibody was studied in the first set of DOE experiments. Because two plates and two QSC preparations were used for the testing, the plate-to-plate and QSC bead's preparation-to-preparation effects were also evaluated (Table 4). Although many factors exhibited

![Figure 6](image-url)
statistical significance, the magnitude of the differences (reported as RSD) was small. Therefore, we concluded that none of the factors tested were practically significant, and the assays were robust within the range of the conditions tested.

3.8.5. DOE 2 for stability of reagents

The second DOE assessed the short-term stability of critical assay reagents. Because two preparations of QSC beads were used in the experiment, the effect of QSC preparations was also evaluated (Table 4). The storage times for the stain solution and FACS buffer were statistically significant for both cell lines but the magnitude of effect (reported as RSD) was small. Thus, no practical significance was observed for the short term stability of the reagents tested in the experiment.

3.8.6. DOE 3 for temperature and exposure to light

The effects of staining temperature and exposure to light were studied in the third DOE. As shown in Table 4, staining temperature and light condition were statistically significant factors, with a consistent and significant decrease in ABC values when the staining procedure was performed at the room temperature. In addition, there was a consistent increase in ABC values when the staining procedure was performed under full room light. However, the magnitude of the differences was smaller than the largest reported intermediate precision for CG1940 or CG8711 and, therefore, although the normal staining procedure requires incubating on ice and in dark, deviation from that procedure is not expected to significantly alter the assay outcome.

Fig. 7. Effect of prolonged storage post thaw and wash for HER-2/neu expression on CG1940 and CG8711. (A) CG1940 post-thaw stability. (B) CG1940 post-wash stability. (C) CG8711 post-thaw stability. (D) CG8711 post-wash stability.

Table 4

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Cell line</th>
<th>Factor</th>
<th>p Value</th>
<th>Statistically significant? (If yes, RSD between conditions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOE1 CG1940 Cells per reaction</td>
<td>0.0773</td>
<td>No</td>
<td>Thaw/stain/wash manipulation</td>
<td>&lt;0.0001 Yes (1.9%)</td>
</tr>
<tr>
<td>DOE1 CG1940 Cells per reaction</td>
<td>0.4184</td>
<td>No</td>
<td>Post-stain acquisition time</td>
<td>0.9548 No</td>
</tr>
<tr>
<td>DOE1 CG1940 Plate</td>
<td>0.0084</td>
<td>Yes (0.6%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE1 CG1940 QSC preparation</td>
<td>0.0417</td>
<td>Yes (0.4%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE1 CG8711 Cells per reaction</td>
<td>0.0345</td>
<td>Yes (0.6%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE1 CG8711 Post-stain acquisition time</td>
<td>&lt;0.0001 Yes (2.4%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE1 CG8711 Plate</td>
<td>0.0004</td>
<td>Yes (1.4%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE1 CG8711 QSC preparation</td>
<td>0.0197</td>
<td>Yes (0.7%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE2 CG1940 FACS buffer age</td>
<td>0.0045</td>
<td>Yes (1.3%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE2 CG1940 Stain solutions age</td>
<td>0.8423</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE2 CG8711 FACS buffer age</td>
<td>&lt;0.0001 Yes (2.5%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE2 CG8711 Stain solutions age</td>
<td>0.7103</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE2 CG8711 QSC preparation</td>
<td>0.9646</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE 3 CG1940 Temperature</td>
<td>&lt;0.0001 Yes (5.0%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE 3 CG1940 Light exposure</td>
<td>&lt;0.0001 Yes (2.8%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE 3 CG8711 Temperature</td>
<td>&lt;0.0001 Yes (5.8%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE 3 CG8711 Light exposure</td>
<td>0.0333 Yes (1.9%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE4 CG1940 Hold time post thaw</td>
<td>0.2894</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE4 CG1940 Hold time pre-stain</td>
<td>0.0003 Yes (1.4%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE4 CG1940 QSC preparation</td>
<td>0.7931</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE4 CG8711 Hold time post thaw</td>
<td>&lt;0.0001 Yes (3.6%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE4 CG8711 Hold time pre-stain</td>
<td>0.3662 No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOE4 CG8711 QSC preparation</td>
<td>0.8616 No</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.8.7. DOE 4 for hold time at various steps

The effect of hold time between the thawing and the first centrifugation steps, and hold time between the second dilution and the staining steps was examined in the last set of DOE experiment. Two preparations of QSC beads were tested in the experiment, so the effect of QSC preparation was also evaluated. Neither of the factors produced practical significant differences between the low and high ends of the hold time (Table 4), and the assay was therefore considered sufficiently robust for these factors.

4. Discussion

Flow cytometry has been used in clinical laboratories for many decades as a diagnostic tool in hematology and hematopathology, and guidelines for assay standardization and validation to ensure the data quality for clinical diagnostic applications are well documented [14–19]. Only certain aspects of the same standardization approaches could be used for our products because, unlike CG1940 and CG8711, clinical specimens typically have significant changes in biomarker expression level and frequency between test (e.g., cells from diseased patients) and control (e.g., healthy patients) samples [20]. For the purposes of lot release and stability studies of a whole cell immunotherapy, the flow cytometer assay can be expected to detect a narrower range of antigen expression levels. Quantification of the biomarker expression level on clinical samples is typically limited to qualitative descriptions of “dim”, “moderate”, and “bright” staining. In more recent years, efforts have been made to translate these descriptive terms to quantitative terms such as molecules of equivalent soluble fluorochrome (MESF) or antibody binding capacity (ABC) through the use of bead-based calibration systems. However, to our knowledge no quantitative flow cytometer assays that measure the biomarker expression at cellular level with high accuracy and precision, and thus serve the purpose of ensuring that the expression levels of samples meet a pre-defined narrow range, have been reported.

The biggest challenge of validating this quantitative flow cytometry assay was the design of the accuracy study because of the lack of established cellular references with known amounts of HER-2/neu expression, as well as the absence of a well-recognized orthogonal quantitative method. Many ideas were considered while searching for scientifically sound approaches to determine assay accuracy. Unlike the assays that determine frequency of expression-positive cells, the distribution of biomarker expression within a cell population does not lend itself to mixing studies wherein HER-2/neu-positive cells are mixed with HER-2/neu-negative cells to demonstrate assay accuracy. In addition, cell surface antigen expression could not be manipulated by independent means (e.g., addition of known amounts of the soluble HER-2/neu protein). Using genetic manipulation to alter the HER-2/neu expression level, such as using anti-sense RNA to inhibit HER-2/neu protein synthesis, was considered but not pursued because it would not produce cells with known specific amounts of HER-2/neu, i.e., it could not routinely produce a known reference standard. Immunological methods such as Western blotting and ELISA of cell lysate had the shortcoming of not being able to distinguish between the cell surface and intracellular HER-2/neu. The Mesoscale Discovery (MSD) electrochemiluminescence detection system (Gaithersburg, MD) was evaluated by loading the whole cells onto the carbon base of the wells and staining with anti-HER-2 antibody. Although the preliminary results are encouraging (data not shown), this system cannot provide population distribution information and thus was only suitable for expression levels and not percent-positive data. Therefore, the Cellomics system which provided quantitative readout of individual cells [10–13] was selected as the orthogonal method to the flow cytometer. The ability of the Cellomics system to serve as a suitable orthogonal method to flow cytometry was verified using several approaches, including QSC beads, differential staining, and the use of multiple different cell lines with known relative amounts of HER-2/neu expression. The correlation between the results generated by the quantitative flow cytometry assay and the Cellomics imaging system demonstrated the suitability of this approach for ascertaining assay accuracy.

The ABC readout for quantifying the HER-2/neu expression level on the cell surface was selected. This method compares the cell surface staining to that on QSC beads that were conjugated with known amounts of goat anti-mouse antibody. This system offers the advantage of using the same antibody fluorochrome conjugate on both the QSC beads and the cell surface and exposing them to the same staining conditions, including pH, ionic strength, and light exposure. Minor variation in antibody conjugation or fluorochrome signal strength would not have a significant impact on final results since this system relies on relative comparison. Because HER-2/neu is typically not expressed profusely on cell surfaces, interference with binding because of the size of fluorochrome-conjugated antibody was not a concern. PE was selected as the fluorochrome for the detection antibody because of its brightness, stability and resistance to self-quenching. The linear titration results of decreasing anti-HER-2/neu antibody confirmed that the use of PE conjugated antibody for this assay was appropriate.

Robustness of the statistical ranges for critical assay steps is usually evaluated in typical validation studies by testing a slightly wider range than the parameters defined in the assay protocol. Both practical and statistical considerations were used to interpret the results from the robustness studies. None of the conditions tested or their interaction caused a practical difference in the ABC values of the sample, although some conditions did have a statistically significant, albeit small, impact. Such factors, including light exposure and antibody incubation time, can be controlled by the procedure. Therefore, this assay was robust within its defined operational ranges. Reproducibility of flow cytometer-based assays was reported as a common concern for comparing results from multiple laboratories [21]. A good instrument calibrator is critical not only for ensuring long-term assay performance but especially for assay transfer to other laboratories. Our results indicated an acceptable level of inter-laboratory variation as compared to reported values from other laboratories [21]. The use of Rainbow beads likely helped to reduce the inter-laboratory difference in instrument settings, which has previously been reported [22,23].

The precision of this assay was evaluated by replication of samples and experiments. Assay precision (the highest RSD is <7%) was within the range of immunological assays and was typical or better than ELISA assays, which have reported RSDs between 10% and 15% [24–26]. In addition to the HER-2/neu assay, we also have developed and validated flow cytometry assays for quantifying the expression level of CD13 on CG1940. The validation conclusion for the CD13 assay was similar to that of the HER-2/neu assay in terms of accuracy, precision, linearity, range, and robustness (data not shown). The successful validation demonstrates that quantitative flow cytometer assays could be a reliable quantitative platform for measuring antigen expression on cell therapy products.

Acknowledgments

The authors are grateful to Bill Hyun at UCSF for help on the Cellomics study, to Debbie Farson, Paul Escarpe, and their group members at Cell Genesys, Inc. for providing cells, and to the Manufacturing group for providing final products of CG1940 and CG8711. We also thank other Cell Genesys colleagues, including Doug Guppill, Sayyou Ohshima, Manjula Chinnappa, Diane DeNagel, Paul Husak, Flavia Borellini, and Junko Aim for providing inputs for...
the early version assay and Cahn To, Christy Cowen, Karen Trujillo, and Charlene Bernard for supporting the assay transfer.

References


