**EVALUATION OF ATUOMATIC CLASS II DESIGNATION FOR**
**clonoSEQ® ASSAY**
**DECISION SUMMARY**

## A. DEN Number:

DEN170080

## B. Purpose for Submission:

De novo request for evaluation of automatic class III designation for the clonoSEQ® Assay

## C. Measurand:

Rearranged IgH (VDJ), IgH (DJ), IgK, and IgL receptor gene sequences, as well as translocated BCL1/IgH (J) and BCL2/IgH (J) sequences.

## D. Type of Test:

Multiplex polymerase chain reaction and next generation sequencing-based *in vitro* diagnostic to evaluate minimal residual disease

## E. Applicant:

Adaptive Biotechnologies Corporation

## F. Proprietary and Established Names:

Adaptive Biotechnologies clonoSEQ Assay

## G. Regulatory Information:

1. Regulation section:

   21 CFR 866.6100

2. Classification:

   Class II (Special Controls)

3. Product code:

   QDC

4. Panel:

   88 − Pathology

## H. Indications For Use:

### 1. Indications for Use:

The clonoSEQ Assay is an in vitro diagnostic that uses multiplex polymerase chain reaction (PCR) and next-generation sequencing (NGS) to identify and quantify rearranged IgH (VDJ), IgH (DJ), IgK, and IgL receptor gene sequences, as well as translocated BCL1/IgH (J) and BCL2/IgH (J) sequences in DNA extracted from bone marrow from patients with B-Cell acute lymphoblastic leukemia (ALL) or multiple myeloma (MM).

The clonoSEQ Assay measures minimal residual disease (MRD) to monitor changes in burden of disease during and after treatment. The test is indicated for use by qualified healthcare professionals in accordance with professional guidelines for clinical decision-making and in conjunction with other clinicopathological features.

The clonoSEQ Assay is a single-site assay performed at Adaptive Biotechnologies Corporation.

### 2. Special conditions for use statement(s)

For prescription use.

For *in vitro* diagnostic use.

### 3. Special instrument requirements:

Illumina NextSeq™ 500 System (qualified by Adaptive Biotechnologies)

## I. Device Description

A description of required equipment, software, reagents, vendors, and storage conditions were provided, and are described in the product labeling. Adaptive Biotechnologies assumes responsibility for the device.

### 1. DNA Extraction

DNA extraction methods have been validated using volumes ranging from 250 µL to 1 mL bone marrow aspirate (BMA)collected in (b) (4) The extraction method isolates DNA by first lysing cells and denaturing proteins after which the DNA is bound to a substrate. Once the DNA is bound, a series of wash steps removes impurities. Following

the wash steps, the DNA is eluted from the substrate. DNA is quantified using a spectrophotometer; the measured DNA concentration is used to add up to 20µg of gDNA to the assay. The MRD test can be performed with 500ng – 20µ g gDNA. Internal controls in the PCR and sequencing steps are used to confirm that sufficient gDNA has been amplified and that amplification was successful.

## 2. Library preparation

Genomic DNA is amplified using locus-specific multiplex PCR using V, D, and J gene primers containing molecular barcodes to amplify IgH (V(D)J), IgH (DJ), IgK, IgL, BCL1/IgH (J), BCL2/IgH (J), and housekeeping gene (HKG) sequences. Reaction specific index barcodes for sample identification are added to the amplified receptor sequences by PCR. Sequencing libraries are prepared by pooling barcoded amplified DNA. qPCR is used to verify the adequacy of the pooled amplified DNA library concentration.

## 3. Sequencing and data analysis

Sequencing is conducted with the Illumina NextSeq$^{TM}$ 500 System. The sequencing process incorporates multiple quality checks. Sequencing output is then processed by the bioinformatics pipeline software as follows:

**a. Flowcell level metrics:**
The analysis pipeline performs quality control (QC) checks on the flowcell data. The pipeline evaluates the percentage of reads that pass the Illumina quality filter (%PF), which must be greater than 70% of reads. The system uses spike-in PhiX templates to evaluate the error rate. The pipeline evaluates the proportion of PhiX reads, which must be greater than 2%, and the associated error rate as computed by the Illumina RTA software, which must be less than 3%.

**b. Demultiplexing and FASTQ generation:**
The pipeline uses Illumina software to demultiplex reads from the instrument output run folder. The analysis pipeline performs a QC check to evaluate whether unexpected barcodes are observed and raises a flag if more than 30,000 reads carry a barcode not specified in the input sample sheet.

**c. Read assignment to receptors:**
The pipeline assigns reads to rearranged receptors for each sample after demultiplexing.

**d. Clonal sequence determination:**
After assigning reads to receptor loci, the pipeline then clusters reads into clonal receptor sequences.

**e. Sample Level QC Checks:**

The pipeline performs a series of sample level QC checks: assessment that sequencing data is sufficient and acceptable based on amplification of sets of internal synthetic controls, assessment that sufficient gDNA is sampled, and a final screen of the calculated values for biologic relevance. One set of internal synthetic controls are evaluated for sufficient read quantity per molecule and read coverage across receptor loci. Another set of internal controls' presence or absence is used to screen for the expected degradation of residual primers. The estimated mass of input gDNA based on an optical density measurement and the estimated number of sampled nucleated cells based on amplification of a set of internal reference genes are used as metrics to check if sufficient material is sampled. The pipeline also checks that the detected number of total and B cells are within a biologically relevant range, and screens for clone sharing by evaluating if sequences are shared across samples that are process together.

**f. Calibrations**

Clonal sequences are assessed for their suitability as ID sequences (to be used for subsequent tracking) by first aggregating highly similar sequences and requiring that the frequency of the sequence is at least 3% as a percentage of all sequences in the locus. The clone must also have a frequency of at least 0.2% of all nucleated cells in the sample and must have sufficient abundance and differentiation from a polyclonal background. Each sequence that is being considered for MRD tracking is compared against a B cell repertoire database and assigned a uniqueness value that, together with its abundance relative to other sequences, is used to assign the sequence to a sensitivity bin which will be used in the estimation of the reported limit of detection (LoD) and limit of quantitation (LoQ).

**g. Tracking:**

When a previous calibration test has identified suitable ID sequences for tracking, they are compared to sequences in the most recent tracking sample in order to assess residual disease. After approximate matching, which allows for mutations in the sample clones as compared to the ID sequences, sequence proportions in the sample are assessed and compared to the LOD and LOQ values. The analysis pipeline then reports whether ID sequences were detected above the LOQ, above LOD but below LOQ, below LOD, or not detected.

**h. Control Materials:**

The following controls are used to measure the success of DNA extraction, PCR amplification, and sequencing:

***i.*** **Synthetic Internal Controls:**

Each sample includes two sets of internal synthetic controls. The controls are panels of synthetic analogues of somatically rearranged B-cell receptor (BCR) immune receptor molecules. The composition of the reference template pools

before and after amplification is measured and used for QC. One set of synthetic templates is added to every pre-amp PCR well as a positive control; these synthetic templates are used to measure primer performance, including identification and correction of amplification bias, and to screen for sufficient sequencing coverage. Another set is added after a step used to remove residual primers.The lack of amplification of these molecules is used to confirm the success of primer removal.

***ii.*** **DNA Extraction Process Controls:**
Each extraction is performed with Positive and Negative Extraction Controls. The Extraction Negative Control is used to confirm lack of contamination during the extraction process. The Extraction Negative Control is subsequently amplified and sequenced in the same fashion as test samples. The Extraction Positive Control is included to assess effectiveness of the extraction process (it is required to be above a pre-set threshold for DNA recovery). If readily available, source material for Extraction Positive Controls is matched to the specimen source type. Exception: The Extraction Positive Control for bone marrow specimens consists of frozen human whole blood.

***iii.*** **PCR Amplification Process Controls:**
Each PCR amplification is performed with an Amplification Positive and Negative Control and subsequently sequenced in the same manner as test samples. The Amplification Positive Control consists of gDNA derived from peripheral blood mononuclear cells and serves as an additional check to confirm successful product amplification. Buffer (1xTE) is used as the Amplification Negative Control.

***iv.*** **Sequencing Process Controls:**
To every sequencing flow cell, two sequencing controls are added. Both a PhiX control purchased from Illumina and a well-characterized amplified library (Sequencing Positive Control) are loaded with test samples.

**i. Result Reporting:**
The pipeline renders results into a PDF-formatted patient report. The report displays any ID sequences identified in the sample that can be used for tracking, with their quantitation and sample-level metrics. For tracking tests, the report includes a result (ID sequences detected above LOD, below LOD, below LoQ, or not detected) and quantitation for the tracked sequences within the most recent sample.

## J. Standards/Guidance Documents Referenced

CLSI guideline EP06-A Evaluation of the Linearity of Quantitative Measurement

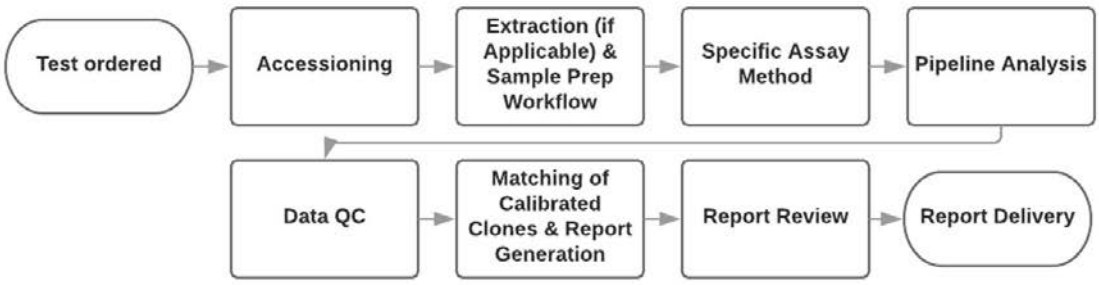
Procedures- A Statistical Approach

## K. Test Principle:

The clonoSEQ Assay is an in vitro diagnostic assay that uses multiplex polymerase chain reaction (PCR) and next-generation sequencing (NGS) to identify and quantify rearranged IgH (VDJ), IgH(DJ), IgK, and IgL receptor gene sequences, as well as translocated BCL1/IgH(J) and BCL2/IgH(J) sequences. The assay also includes primers that amplify specific genomic regions present as diploid copies in normal genomic DNA (gDNA) to allow determination of total nucleated cell content.

Testing begins with gDNA extracted from bone marrow (Figure 1). Extracted gDNA quality is assessed and rearranged immune receptors are amplified using a multiplex PCR. Reaction specific index barcode sequences for sample identification are added to the amplified receptor sequences by PCR. Sequencing libraries are prepared from barcoded amplified DNA, which are then sequenced by synthesis using NGS. Raw sequence data are uploaded from the sequencing instrument to the Adaptive analysis pipeline. These sequence data are analyzed in a multi-step process: first, a sample's sequence data are identified using the sample index sequences. Next, data are processed using a proprietary algorithm with in-line controls to remove amplification bias. When the clonoSEQ Clonality (ID) assessment is conducted, the immune repertoire of the sample is checked for the presence of DNA sequences specific to "dominant" clone(s) consistent with the presence of a lymphoid malignancy. Each sequence that is being considered for MRD tracking is compared against a generic B cell repertoire database and assigned a uniqueness value based on the observed or estimated abundance of that sequence in the database. The uniqueness value is used to assign the sequence to a sensitivity bin which will be used in the estimation of the reported limit of detection (LoD) and limit of quantitation (LoQ) on the patient report. During clonoSEQ Tracking (MRD) assessment, the complete immunoglobulin receptor repertoire is again assessed, and the previously identified dominant clonotype sequence(s) are detected and quantified to determine the sample MRD level. The clonoSEQ assay MRD assessment measures MRD disease in a biologic sample.

**Figure 1: clonoSEQ Assay Workflow**




Following completion of these data processing steps, a report is issued. A Clonality (ID) report indicates the presence of dominant sequences residing within a presumed

malignant lymphocyte clonal population, as identified in the baseline (diagnostic or high disease burden) sample from a patient. After one or more dominant sequence(s) have been identified in a baseline sample, subsequent samples from the same patient can be assessed for MRD, after which a Tracking (MRD) report is generated. The MRD is expressed as a frequency that quantifies the level of residual disease (based on the number of remaining copies of the initially dominant sequence(s) relative to the total number of nucleated cells in the sample).

## L. Performance Characteristics:

1. Analytical Performance

*a. Functional Equivalence of gDNA Blends to Cellular Blends*

Clinical specimens from 23 patients with MM and 21 patients with ALL were used for precision, quantitative accuracy and linearity studies. An additional 22 specimens from patients with other lymphoid malignancies were used to supplement the analytical sensitivity studies. Sample types included bone marrow aspirate (BMA), bone marrow mononuclear cells (BMMCs), CD138+ bone marrow cells, peripheral blood, and peripheral blood mononuclear cells (PBMCs). gDNA was isolated from these clinical samples and blended with gDNA isolated from bone marrow to contrive specific MRD levels for the analytical studies.

In order to evaluate whether the blended DNA from clinical specimen that are described above were suitable for specific analytical studies, a study was performed to evaluate whether MRD estimates from blended gDNA (i.e., gDNA first isolated from specimens and then blended with gDNA from peripheral blood mononuclear cells to create specific MRD levels) were equivalent to MRD estimates from gDNA isolated from blended specimens (i.e., specimens were blended with peripheral blood mononuclear cells at known concentrations prior to gDNA isolation). The accuracy and linearity of the assay results were assessed across 11 MRD frequency levels, ranging from $3.3x10^{-7}$ to $3.0x10^{-3}$, for both blended gDNA and DNA extracted from blended cells. These dilutions included levels below the LoD and spanned the range of reportable MRD levels. The MRD estimates on gDNA blends were comparable to the MRD estimates of the blended cells they were intended to mimic. Therefore, the blended gDNA from the clinical samples was determined to be functionally equivalent to clinical specimens for use in specific analytical studies.

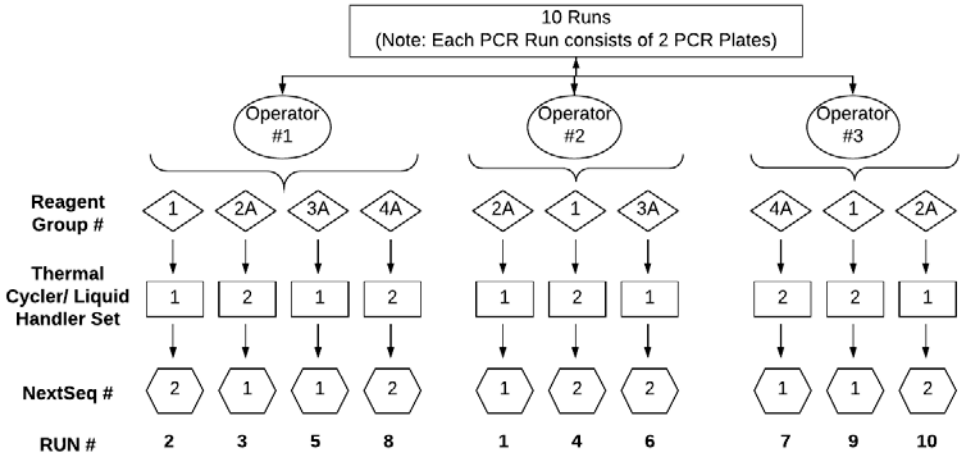
*b. Precision/Reproducibility Studies*

Precision studies tested gDNA extracted from clinical specimens from 23 patients with MM and 21 patients with ALL. The gDNA from these specimens was used to contrive specific MRD levels by pooling and blending them into gDNA extracted from the BMA of healthy donors. The study included three DNA inputs (500ng, 2µg,

20µg) and six MRD levels were tested at each DNA input for each patient sample. The studies were designed to test the MRD levels of $2.8x10^{-5}$, $8.0x10^{-5}$, $2.8x10^{-4}$, $8.0x10^{-4}$, $2.8x10^{-3}$ and $8.0x10^{-3}$ at 500 ng DNA input; $7.0X10^{-6}$, $2.0x10^{-5}$, $7.0x10^{-5}$, $2.0x10^{-4}$, $7.0x10^{-4}$ and 2.0x10-$^{3}$ at 2 µg DNA input; and 7.0 $x10^{-7}$, $2.0x10^{-6}$, $7.0x10^{-6}$, $2.0x10^{-5}$, $7.0x10^{-5}$ and $2.0x10^{-4}$ at 20 µg DNA input. These frequencies correspond to an estimated 2.14, 6.13, 21.44, 61.26, 214.40, and 612.56 malignant cells tested at each DNA dilution.

The precision study used a main effects screening design over 21 calendar days. This study used 10 runs, with 2 PCR plates in each run, using 3 operator sets, 4 reagent lots, and 4 instrument sets (2 thermal cycler/liquid handlers and 2 NextSeq instruments). The study design is summarized in Figure 2.

**Figure 2: Precision Study Design Schematic**




Each run of the assay tested 18 combinations of DNA input and MRD frequency in duplicate. In all, 360 contrived samples were tested. Of these, one plate with 18 samples was invalid due to sample QC failures; the plate-level failure rate was therefore 1 / 20 = 0.05. An additional two contrived samples (88 MRD measurements) failed sample QC due to insufficient sequencing coverage. While normal operating procedures permit re-sequencing, for this analysis these 2 samples were classified as failures. The analysis used the remaining 340 contrived samples with up to 44 MRD measurements per sample, for a total of 14,744 MRD measurements.

The precision of the clonoSEQ assay is largely dependent upon the number of malignant cells that are being evaluated, rather than the actual MRD frequency. Consequently, the same MRD frequency is expected to have lower precision at lower DNA inputs. For this study, precision estimates were first calculated based on the MRD frequency per DNA input, followed by estimates of imprecision of the absolute number of malignant cells detected per reaction.

i. Precision of MRD Frequency for MM and ALL

Precision analysis, including variation from instrument set, operator, processing day, processing run, and reagent lot, is reported as %CV for each tested MRD frequency at each DNA input. The analysis was done separately for MM and ALL, and the data are summarized in Tables 1 and 2. Precision ranged from 29.6% to 70.0% CV for MM and 25.9% to 76.9% CV for ALL.

In these tables, MRD frequency range refers to the central 95% range of MRD estimates that were observed across all of the patient samples tested at each DNA input and frequency condition. These data were used to define the 95% confidence intervals that are used in the patient reports.

**Table 1: Precision of the clonoSEQ Assay in MM Samples**

| DNA Input | MRD Frequency | Measurements | Patients* | %CV | Frequency Range (95% CI) |
|---|---|---|---|---|---|
| 500 ng | $3.5x10^{-5}$ | 378 | 21 | 70.0 | 0 - $9.0x10^{-5}$ |
| | $8.9x10^{-5}$ | 378 | 21 | 52.6 | $1.9x10^{-5}$ - $2.0x10^{-4}$ |
| | $3.0x10^{-4}$ | 414 | 23 | 37.6 | $1.2x10^{-4}$ - $5.7x10^{-4}$ |
| | $8.4x10^{-4}$ | 460 | 23 | 34.8 | $3.9x10^{-4}$ - $1.6x10^{-3}$ |
| | $2.7x10^{-3}$ | 460 | 23 | 32.3 | $1.4x10^{-3}$ - $5.1x10^{-3}$ |
| | $6.5x10^{-3}$ | 460 | 23 | 30.1 | $3.6x10^{-3}$ - $1.2x10^{-2}$ |
| 2 µg | $8.1x10^{-6}$ | 378 | 21 | 69.2 | 0 - $2.1x10^{-5}$ |
| | $2.3x10^{-5}$ | 378 | 21 | 52.1 | $5.7x10^{-6}$ - $5.4x10^{-5}$ |
| | $7.8x10^{-5}$ | 414 | 23 | 38.8 | $3.1x10^{-5}$ - $1.5x10^{-4}$ |
| | $2.1x10^{-4}$ | 460 | 23 | 34.4 | $1.0x10^{-4}$ - $4.0x10^{-4}$ |
| | $7.4x10^{-4}$ | 460 | 23 | 32.4 | $3.7x10^{-4}$ - $1.4x10^{-3}$ |
| | $2.0x10^{-3}$ | 460 | 23 | 30.2 | $1.1x10^{-3}$ - $3.8x10^{-3}$ |
| 20 µg | $8.6x10^{-7}$ | 378 | 21 | 65.6 | 0 - $2.1x10^{-6}$ |
| | $2.3x10^{-6}$ | 378 | 21 | 51.7 | $5.1x10^{-7}$ - $5.0x10^{-6}$ |
| | $8.2x10^{-6}$ | 414 | 23 | 37.2 | $3.5x10^{-6}$ - $1.5x10^{-5}$ |
| | $2.3x10^{-5}$ | 460 | 23 | 33.3 | $1.2x10^{-5}$ - $4.5x10^{-5}$ |
| | $7.6x10^{-5}$ | 460 | 23 | 31.1 | $4.1x10^{-5}$ - $1.4x10^{-4}$ |
| | $2.1x10^{-4}$ | 414 | 23 | 29.5 | $1.2x10^{-4}$ - $3.7x10^{-4}$ |

* Note: Some contrived samples included a subset of patient samples

**Table 2: Precision of the clonoSEQ Assay in ALL Samples**

| DNA Input | MRD Frequency | Measurements | Patients | %CV | Frequency range (95% CI) |
|---|---|---|---|---|---|
| 500 ng | $3.2x10^{-5}$ | 378 | 21 | 76.9 | 0 - $9.3x10^{-5}$ |
| | $8.6x10^{-5}$ | 378 | 21 | 54.1 | $1.0x10^{-5}$ - $1.9x10^{-4}$ |
| | $2.8x10^{-4}$ | 378 | 21 | 33.0 | $1.0x10^{-4}$ - $4.8x10^{-4}$ |
| | $7.6x10^{-4}$ | 420 | 21 | 29.3 | $3.4x10^{-4}$ - $1.2x10^{-3}$ |
| | $2.5x10^{-3}$ | 420 | 21 | 26.5 | $1.0x10^{-3}$ - $4.0x10^{-3}$ |
| | $6.0x10^{-3}$ | 420 | 21 | 25.9 | $2.2x10^{-3}$ - $9.8x10^{-3}$ |
| 2 µg | $7.6x10^{-6}$ | 378 | 21 | 74.1 | 0 - $2.0x10^{-5}$ |
| | $2.1x10^{-5}$ | 378 | 21 | 47.4 | $3.8x10^{-6}$ - $4.0x10^{-5}$ |
| | $6.8x10^{-5}$ | 378 | 21 | 33.3 | $2.2x10^{-5}$ - $1.1x10^{-4}$ |
| | $1.9x10^{-4}$ | 420 | 21 | 29.1 | $7.0x10^{-5}$ - $3.1x10^{-4}$ |
| | $6.8x10^{-4}$ | 420 | 21 | 27.1 | $2.5x10^{-4}$ - $1.1x10^{-3}$ |
| | $1.9x10^{-3}$ | 420 | 21 | 25.9 | $6.9x10^{-4}$ - $3.1x10^{-3}$ |
| 20 µg | $8.5x10^{-7}$ | 378 | 21 | 74.8 | 0 - $2.2x10^{-6}$ |
| | $2.3x10^{-6}$ | 378 | 21 | 49.8 | $4.0x10^{-7}$ - $4.7x10^{-6}$ |
| | $7.5x10^{-6}$ | 378 | 21 | 34.5 | $2.6x10^{-6}$ - $1.3x10^{-5}$ |
| | $2.1x10^{-5}$ | 420 | 21 | 27.9 | $8.5x10^{-6}$ - $3.4x10^{-5}$ |
| | $6.9x10^{-5}$ | 420 | 21 | 26.7 | $2.5x10^{-5}$ - $1.1x10^{-4}$ |
| | $2.0x10^{-4}$ | 378 | 21 | 26.5 | $7.1x10^{-5}$ - $3.3x10^{-4}$ |

ii. Precision of Malignant Cells Detected

The precision for detection of malignant cells detected was evaluated across a range of tested malignant cells (2.14 – 612.56). For this analysis, the results from all of the DNA inputs across MM and ALL were pooled into a single analysis that is summarized in Table 3. As expected, the precision was primarily influenced by the number of cells being evaluated. Precision ranged from %CV = 72% at 2.14 cells to %CV = 21% at 612.56 cells. The tested factors (Operator, Instrument Sets, Reagent Lots, Day, Run) minimally contributed to variability, with attributable %CV ranging from 0% to 3% (Table 3). The majority of the observed variation is due to the residual variability. Residual variability refers to the amount of variability remaining after accounting for the variability explained by the tested factors.

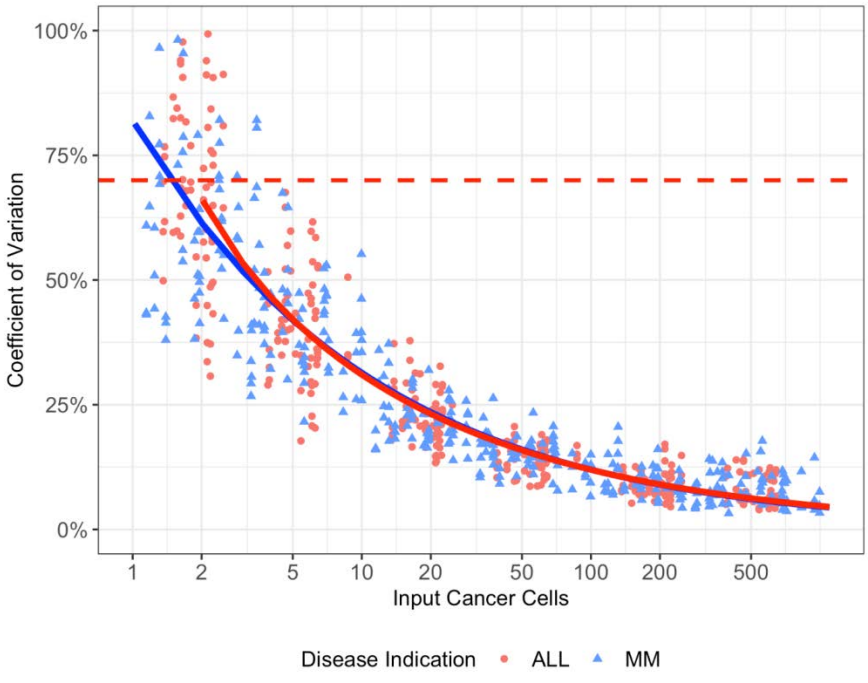
**Table 3: Summary of the clonoSEQ Assay Precision of Malignant Cells Detected**

| | | %CV Attributed to Each Variable at Cell Inputs* | | | | | |
|---|---|---|---|---|---|---|---|
| | **# of Input Cancer Cells** | **2.14** | **6.13** | **21.44** | **61.26** | **214.4** | **612.56** |
| | **Instrument Set** | 0% | 1% | 0% | 1% | 1% | 0% |
| | **Operator** | 0% | 1% | 0% | 2% | 0% | 0% |
| | **Processing Day** | 0% | 0% | 1% | 0% | 0% | 3% |
| | **Processing Run** | 0% | 0% | 1% | 0% | 0% | 0% |
| **Lot-to-Lot Variability** | **Reagent Lot** | 2% | 0% | 0% | 2% | 2% | 1% |
| | **Residual Variability** | 72% | 51% | 30% | 24% | 20% | 19% |
| **N** | **Total MRD Measurements** | 2268 | 2268 | 2376 | 2640 | 2640 | 2552 |

* These values were aggregated across diseases (ALL and MM) and total DNA input levels

The precision for each sample at each tested condition across all DNA inputs is summarized in a Sadler's precision profile (Figure 3). The Sadler's precision profile visualizes the relationship between the number of sampled malignant cells and precision, as measured by %CV. This analysis demonstrates that the precision of the clonoSEQ Assay is largely dependent on the number of malignant cells that are being evaluated by the assay.

**Figure 3: Sadler's Precision Profile (Coefficient of Variation) of the clonoSEQ Assay as a Function of Input Cancer Cells**



*c. DNA Extraction Reproducibility*

Precision studies (described above) were conducted with extracted, pooled and aliquoted genomic DNA from patient specimens. In order to determine the impact of the DNA extraction step on the reproducibility of the clonoSeq assay, two reproducibility studies were conducted beginning with bone marrow aspirates. For the first study, BMA from 3 heathy subjects were extracted using 250 µL or 500 µL of BMA. Two (2) replicate aliquots of each sample were extracted in each of 4 separate extraction runs using 2 operators, 2 instruments, and 2 extraction reagent lots. DNA yields from the 250 µL input ranged from 9.6 µg to 21.7 µg of DNA. DNA yields from the 500 µL input of BMA ranged from 18.7 µg to 45.9 µg of DNA. Neither operator, instrument, reagent lot, or extraction run had a significant effect on gDNA yields. All samples had acceptable A260/A280 measurements within the required range of 1.7 to 2.3. The sequencing data were assessed for the impact of the different extraction conditions on the ratio of rearranged B cell receptors to the total number of cells as measured by the clonoSEQ assay and all samples passed pre-established acceptance criteria of ± 30% MRD frequency.

A second study was conducted to evaluate DNA extraction on MRD estimation. Samples were created by diluting bone marrow mononuclear cells (BMMCs) from 3 ALL and 3 MM patients with BMA from healthy subjects to a concentration of ~50 cells per disease sample per 1 million BMA cells. The resulting simulated MRD levels ranged from $4.9 \times 10^{-6}$ to $1.5 \times 10^{-4}$ malignant cells. Four (4) extraction runs were performed on 500 µL BMA and on 1.0 mL BMA, each in duplicate using 2 operators, 2 instruments, and 2 lots. The study met the pre-established acceptance criteria of ± 30% MRD frequency between different operators, instrument sets, and reagent lots.

*d. Precision of Sequence/Nucleotide Base Calls*

The clonoSEQ assay reports out the sequence information for dominant clones which can then be tracked in future measurements. Therefore, the repeatability of sequences generated by of the clonoSEQ Assay was assessed using a two-step process. First, ID samples from 72 lymphoid malignancy samples and 9 cell lines were processed to determine the baseline calibrating clonotype nucleotide sequences. Next, 20 replicates of the samples were run at disease inputs of ~2 to 600 malignant cells across four DNA inputs (10ng, 500ng, 2µg, 20µg). The replicates were tested using 3 operators, 2 instrument sets, and 4 reagent lots. These data were used to assess the observed rate of agreement between the nucleotide sequences chosen for tracking and the nucleotide sequences observed in contrived samples from the same biological specimens.

For each calibrating clonotype sequence in an ID sample, all sequences in the corresponding MRD samples within N bp were included for assessment of overall percent agreement (OPA), where N is defined for each sequence as the number of allowable mutations determined during specimen characterization by the calibration algorithm. N is chosen to capture somatic variation among B cells from the same

clonal lineage without incorrectly grouping sequences from different clonal lineages. Once this population was established, the OPA between the original calibrating clonotype sequence and the sequences identified in the MRD assessment was calculated.

Table 4 reports the number of nucleotides assessed, the OPA, the lower and upper 95% confidence limits, and the OPA restated in the same terms as a Phred quality score (i.e., -10 x log10 disagreement rate). This test assessed approximately 442.5 million nucleotides for sequence agreement, with an overall disagreement rate of approximately 3.5 parts per 100,000 (corresponding to a Phred score of about 44.5; in typical NGS applications a Phred score of 30 or higher constitutes a high-quality base call).

**Table 4: Summary of Sequence Agreement Metrics**

| Allowed Mutations (N) | Nucleotides Assessed | OPA | LCL | UCL | Phred |
|---|---|---|---|---|---|
| **1** | 135,025,044 | 99.9968 | 99.9967 | 99.9969 | 44.9 |
| **2** | 57,248,770 | 99.9965 | 99.9964 | 99.9967 | 44.6 |
| **3** | 151,018,837 | 99.9965 | 99.9965 | 99.9966 | 44.6 |
| **4** | 82,780,612 | 99.9960 | 99.9959 | 99.9962 | 44.0 |
| **5** | 13,918,166 | 99.9966 | 99.9963 | 99.9969 | 44.6 |
| **6** | 2,587,014 | 99.9961 | 99.9953 | 99.9968 | 44.1 |

The data support that the sequences reported by clonoSEQ assay are reproducible with low sequencing error rates. This demonstrates that the sequences identified in the dominant clones during the clonoSEQ Clonality ID assessment can be reliably identified in the clonoSEQ MRD tracking assessment.

*a. Amplification Bias by Clonotype*

Two types of studies were executed to assess amplification bias. One study used a comprehensive panel of synthetic double-stranded molecules representing rearrangements of the targeted exons, while the other used clinical samples. Data from amplification of the synthetic templates demonstrate that the clonoSEQ Assay amplifies the targeted exon segments efficiently and consistently with nominal bias. These conclusions were supported by data from clinical samples, which show that patients who carry certain exons in their malignant clonotypes do not have biased precision profiles.

*e. Linearity*

i. Linearity using Cell Lines

Linearity of the clonoSEQ Assay using 3 ALL cell lines (SUP-B15, GM14952,

and GM20930) and 3 MM cell lines (IM9, U266, and L363) was evaluated by blending cell line gDNA with gDNA from healthy subjects, using DNA inputs of 200 ng, 2 µg and 20 µg gDNA, and tested at zero and across 11 MRD frequencies at each DNA input. The frequency range of $6.0x10^{-5}$ to 1 was tested at the 200 ng DNA input. The frequency range of $6.5x10^{-6}$ to 1 was tested at the 2 µg DNA input. The frequency range of $6.6x10^{-7}$ to 0.1 was tested at 20 µg DNA input. The linear range of the assay was determined by finding the input range where the maximum deviation from linearity (based on a quadratic or cubic fit to the data) was less than 5%.

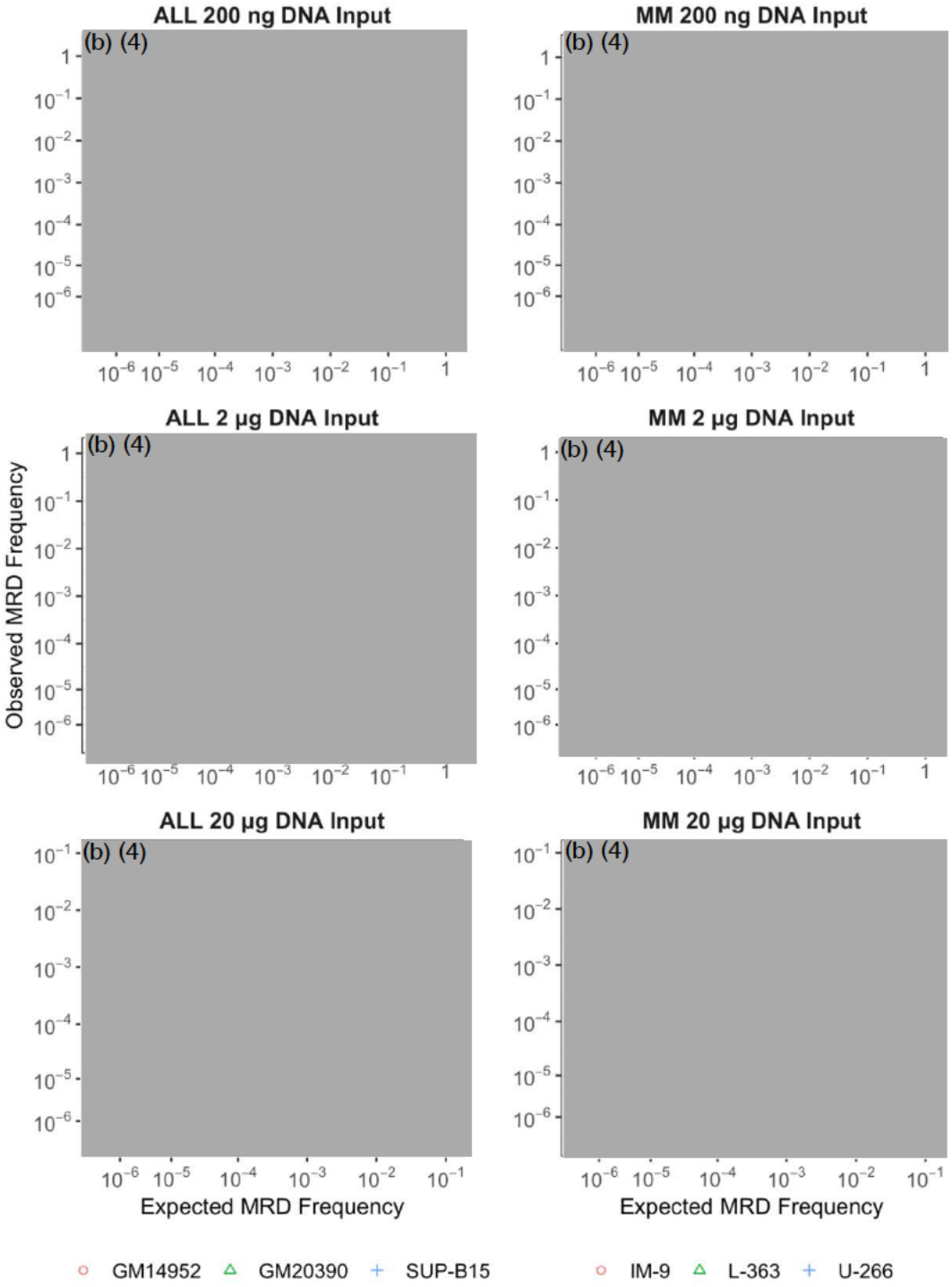
Linearity was established for each sample input and disease type across the entire tested range (Table 5 and Figure). This study demonstrated linearity of MRD frequencies across several orders of magnitude for each condition tested.

**Table 5: Linearity of the clonoSEQ assay using cell lines**

| Disease Indication | Input DNA | Tested Range | Linear Range | Average Slope (Range)* | Average Intercept (Range)* |
|---|---|---|---|---|---|
| ALL cell lines | 200ng | 0 to 100% | 0% to 100% | 1.01 (0.98 to 1.05) | -0.11 (-0.31 to 0.16) |
| | 2µg | 0 to 100% | $3 \times 10^{-5}$ to 30% | 0.98 (0.98 to 1.00) | -0.17 (-0.39 to -0.04) |
| | 20µg | 0 to 10% | 0% to 10% | 0.95 (0.88 to 0.99) | -0.26 (-0.63 to -0.05) |
| MM cell lines | 200ng | 0 to 100% | 0% to 100% | 1.03 (0.98 to 1.11) | -0.07 (-0.08 to -0.04) |
| | 2µg | 0 to 100% | $9.8 \times 10^{-6}$ to 30% | 1.02 (0.98 to 1.06) | -0.03 (-0.08 to 0.02) |
| | 20µg | 0 to 10% | 0% to 10% | 0.98 (0.96 to 0.99) | -0.15 (-0.22 to -0.04) |

*Range represents the range of values across the three cell lines tested.

**Figure 4: Linearity of the clonoSEQ Assay. Expected (x-axis) and Observed (y-axis) MRD Frequency of Six Cell Lines.**




ii. Linearity using Clinical Specimens

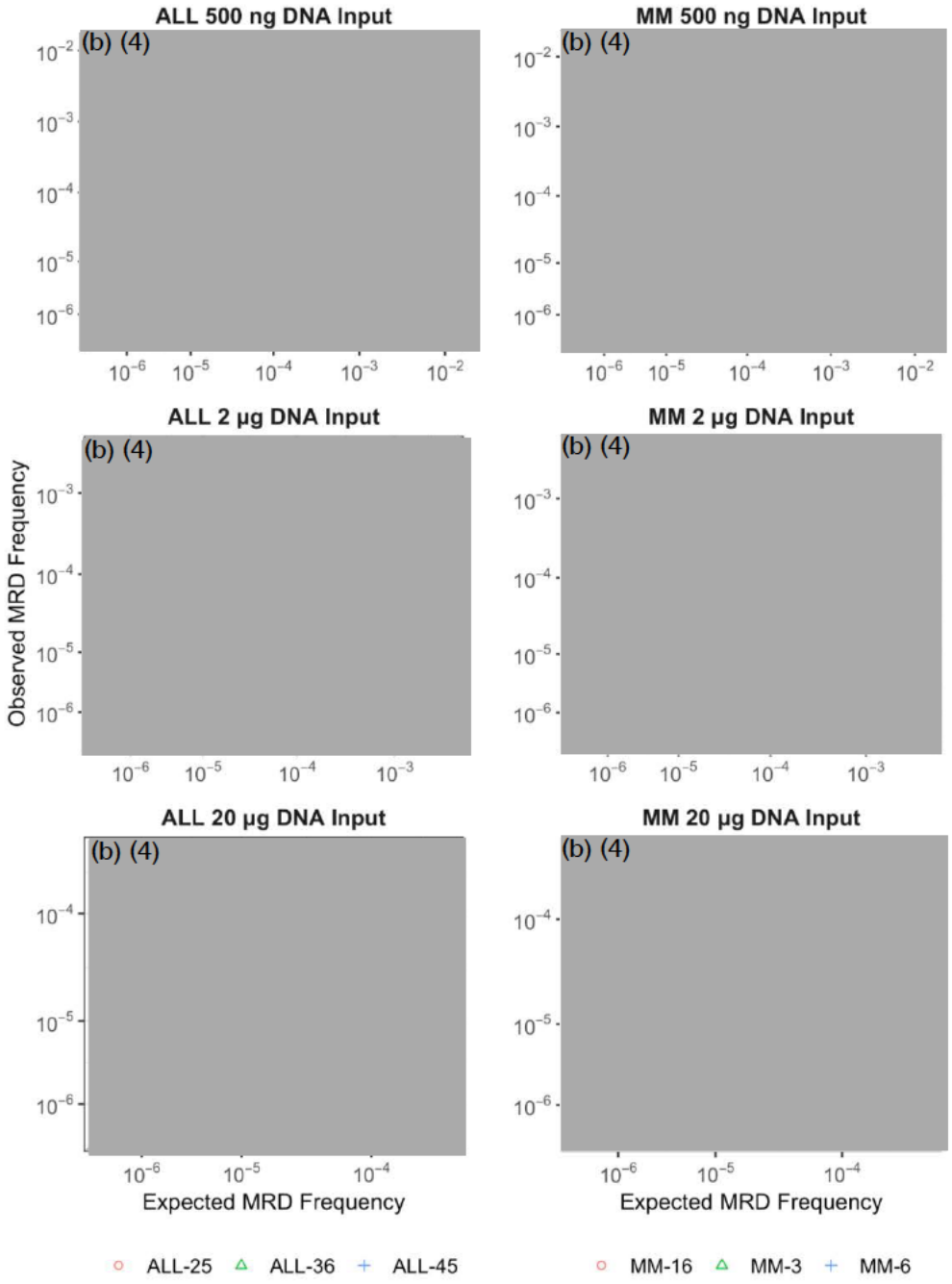
The precision study described in section (L)(1)(b) evaluated blended gDNA extracted from 23 MM and 21 ALL specimens at three DNA inputs and 6 MRD

frequencies per DNA input. These data were re-analyzed to confirm linearity at the lower frequency range of the assay. The linear range of the assay was determined by finding the input range where the maximum deviation from linearity (based on a quadratic or cubic fit to the data) was less than 5%. Results are summarized in Table 6. The slopes and intercepts are reported as the average and range of values across all clinical specimens that were tested at each DNA input by disease indication. This study demonstrated linearity across a wide range of MRD frequencies for each condition tested using clinical specimens. Results from 3 representative specimens for each ALL and MM are visualized in Figure 5.

**Table 6: Linearity of clonoSEQ assay using clinical specimens**

| | | | **Combined Analysis** | | | **Summary of individual patient analyses** | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Disease Indication** | **Input DNA** | **Tested Range** | **Linear Range** | **Slope** | **Intercept** | **Average Slope** | **Average Intercept** | **Slope Range** | **Intercept Range** |
| ALL | 500 ng | $2.8x10^{-5}$ to $8x10^{-3}$ | $2.8x10^{-5}$ to $8x10^{-3}$ | 0.948 | -0.214 | 0.948 | -0.214 | 0.853 to 1.073 | -0.461 to 0.018 |
| | 2 µg | $7x10^{-6}$ to $2x10^{-3}$ | $7x10^{-6}$ to $2x10^{-3}$ | 0.985 | -0.074 | 0.985 | -0.074 | 0.909 to 1.076 | -0.757 to 0.233 |
| | 20 µg | $7x10^{-7}$ to $2x10^{-4}$ | $7x10^{-7}$ to $2x10^{-4}$ | 0.978 | -0.101 | 0.978 | -0.101 | 0.859 to 1.029 | -1.018 to 0.336 |
| MM | 500 ng | $2.8x10^{-5}$ to $8x10^{-3}$ | $2.8x10^{-5}$ to $8x10^{-3}$ | 0.962 | -0.143 | 0.956 | -0.158 | 0.853 to 1.148 | -0.462 to 0.183 |
| | 2 µg | $7x10^{-6}$ to $2x10^{-3}$ | $7x10^{-6}$ to $2x10^{-3}$ | 0.986 | -0.040 | 0.983 | -0.049 | 0.924 to 1.068 | -0.341 to 0.246 |
| | 20 µg | $7x10^{-7}$ to $2x10^{-4}$ | $7x10^{-7}$ to $2x10^{-4}$ | 0.985 | -0.034 | 0.981 | -0.052 | 0.933 to 1.075 | -0.419 to 0.611 |

**Figure 5: Linearity of the clonoSEQ Assay. The Expected (x-axis) and Observed (y-axis) MRD Frequency of Six Representative Clinical Samples.**




*f. Analytical Sensitivity*

i. Limit of Blank (LoB)

The LoB was determined by measuring the specificity of a patient's trackable immunoglobulin (Ig) sequences. These sequences were identified from samples

from 66 patients diagnosed with a lymphoid malignancy (23 MM, 21 ALL, and 22 other malignancy). The LoB was determined by searching for the presence and abundance of these trackable sequences in healthy bone marrow samples. The 95$^{th}$ percentile of sample MRD frequencies for these trackable sequences was zero at 500 ng and 20 µg gDNA input from DNA extracted from healthy bone marrow. Therefore, the LoB was zero, demonstrating that trackable Ig sequences are highly patient-specific.

ii. Limit of Detection (LoD)/Limit of Quantitation (LoQ)

The LoD and LoQ were determined by blending gDNA extracted from 66 specimens from patients with lymphoid malignancies (23 MM, 21 ALL, and 22 with other malignancy) into 500 ng and 20 µg of gDNA extracted from bone marrow of healthy subjects. A dilution series of 22.97, 10.72, 4.59, 2.14, and 0.94 malignant cell equivalents for each patient was made for each DNA input level. Each sample was tested in duplicate for each of four reagent lots, resulting in 8 results for each of the 66 samples at each dilution condition. A probit approach was used to determine the LoD to be 1.903 malignant cells (95% CI; 1.75 – 2.07) based on the combined data from both DNA inputs (Table 7). The LoQ was defined as the lowest absolute number of malignant cells whose frequency can be quantitatively determined with an accuracy of 70% relative total error. The LoQ was found to be 2.390 malignant cells (95% CI; 1.90 – 9.14) (Table 7).

**Table 7: LoD/LoQ by MRD Cell Counts and MRD Frequency**

| Measure | Malignant Cells* | 500 ng DNA Input Frequency (95% CI) | 20 µg DNA Input Frequency (95% CI) |
|---|---|---|---|
| LoD | 1.903 (1.75 - 2.07) | $2.26x10^{-5}$ ($2.01x10^{-5}$ - $2.53x10^{-5}$) | $6.77x10^{-7}$ ($6.02x10^{-7}$ - $7.61x10^{-7}$) |
| LoQ | 2.390 (1.90 - 9.14) | $2.39 \times 10^{-5}$ ($2.26x10^{-5}$ - $7.01x10^{-5}$) | $1.76x10^{-6}$ ($6.77x10^{-7}$ - $4.09x10^{-6}$) |

*Calculated from samples with 500 ng and 20 µg of DNA input.

The clonoSEQ Assay can use a range of DNA inputs from 500 ng to 20 µg of DNA. The LoD/LoQ by MRD frequency will vary based on the DNA input and the total nucleated cells that are evaluated by the assay. The estimated LoD/LoQ at 500 ng and 20 µg of DNA input are shown in Table 7.

To confirm the LoD and LoQ, 1.903 and 2.390 malignant cell equivalents were spiked into 200ng, 500ng, 1µg, 2µg, 5µg, 10µg, 20µg, and 40µg of gDNA extracted from bone marrow of healthy subjects. The results showed that the LoD and LoQ of malignant cells detected remained consistent across all DNA input levels.

g. *Analytical Specificity*

i. Interfering Substances

Testing was performed to characterize the effects of five endogenous (Table 8) and three exogenous (Table 9) substances at the concentrations listed below. The potential exogenous and endogenous substances were spiked separately into 250 µL aliquots of bone marrow from 4 different donors. Each condition was replicated for a total of 8 times (4 donors x 2 replicates) and all conditions passed the pre-specified MRD frequency equivalence margin of ± 30%. This study concluded that MRD results were not substantially influenced by the presence of the tested interfering substances.

**Table 8: Endogenous Interfering Substances Tested**

| Substance Name | Reference Level (Low) | Recommended Concentration (High) | Acceptance Criteria: Pass/Fail |
|---|---|---|---|
| Bilirubin conjugated | 3.4 µmol/L | 342 µmol/L | Pass |
| Bilirubin unconjugated* | 21 µmol/L | 342 µmol/L | Pass |
| Hemoglobin | 1 g/L | 2 g/L | Pass |
| Cholesterol* | 5.2 mmol/L | 13 mmol/L | Pass |
| Triglycerides | 3.7 mmol/L | 37 mmol/L | Pass |

(b) (4) was used as solvent to resuspend bilirubin (unconjugated) and cholesterol.

**Table 9: Exogenous Interfering Substances Tested**

| Substance Name | Concentration (Low) | Concentration (High) | Acceptance Criteria: Pass/Fail |
|---|---|---|---|
| (b) (4) | 1.8 mg/mL | 3.6 mg/mL | Pass |
| (b) (4) | 15 USP U/mL | 30 USP U/mL | Pass |
| (b) (4) | N/A | N/A | Pass |

* The BMA samples were shipped to Adaptive containing 1.8 mg/ml (b) (4) ("Low" concentration) for anti-coagulation purposes. Additional (b) (4) was spiked in to achieve the High level

(b) (4) inhibition was tested at a single spiked-in volume (2.5 µl).

ii. Cross-Contamination/Sample Carryover

The assessment of cross-contamination included two studies: one study to measure contamination of ID samples during automated DNA extraction of BMA and BMMC and one study to measure contamination of gDNA from MRD samples during PCR, library pooling, and sequencing with the clonoSEQ Assay.

Cross-contamination of ID samples during automated DNA extraction was assessed using a panel of 3 ALL cancer cell lines and 3 MM cell lines each spiked to 10% of total cells in a BMA pool of 2 healthy subjects or a bone marrow mononuclear cell (BMMC) pool of 4 healthy subjects. PBS (blank) samples were

included in this study. Samples were evaluated as to whether they correctly calibrated. There were no false calibrations for run-to-run with 0 of 44 BMA and 0 of 44 BMMC false calibrations. There was 1 false calibration for the well-to-well study with 1 of 44 BMA and 0 of 44 BMMC samples falsely calibrating. The falsely calibrated sequence was found in a PBS sample with 83 total templates and the sequence was not associated with any of the 6 cell lines. The PBS sample provides a sensitive test for contamination since there was no background DNA and a contamination of 83 templates would not be expected to cause false calibration of a clinical specimen.

Cross contamination of incorrectly calling samples MRD positive was assessed using gDNA from peripheral blood from healthy subjects as MRD-negative specimens and blends of cell line gDNA and gDNA from peripheral blood of healthy subjects spiked to a concentration of 5%. The 5% level was used to simulate a patient with clinical relapse. This study evaluated for the presence of a clonal sequence and molecular barcode simultaneously. There were no run-to-run contamination events observed (0 of 36). Well-to-well cross contamination was observed in 8 of 712 comparisons; this was likely caused by contamination of a primer barcode plate sourced from a vendor. All contamination events were below $4x10^{-6}$. This low level of contamination is unimpactful to a final MRD call because tracked clonotype sequences are highly specific to each patient, so this level of contamination between samples from different patients would not affect the reported MRD result. Cross contamination between samples from the same patient is prevented by process controls that disallow co-processing of samples from the same patient.

*h. Traceability (controls, calibrators, or methods)*

The clonoSEQ Assay is not traceable to any known standard. Controls and quality metrics are described in the device description section.

*i. Reagent Stability*

i. <u>In-Use Stability</u>

An in-use stability study was executed to determine stability needs of the clonoSEQ Assay for reaction mixes and intermediate steps. The following critical steps were evaluated: pre-amp and PCR primer mix stability, master mix stability, complete reaction stability and process pause stability. (b) (4) gDNA was tested using seven replicates for all conditions tested. Acceptance criteria were based the on sequencing results meeting all QC metrics and all of the conditions tested met the pre-specified acceptance criteria and the clonoSEQ Assay in-use stability needs.

ii. <u>Real Time Stability of Pre-Amp and PCR Mixes</u>

The real-time reagent stability studies used the primer QC processes to assess primer performance and determine primer stability. The primer QC process uses a

set of synthetic double-stranded molecules representing rearrangements of the targeted exons to determine whether each manufactured lot of pre-amp PCR primers and PCR primers are performing within specification. The priming sites on synthetic molecules are identical to biologic priming sites on targeted exons. Data from these molecules were analyzed and assessed for the ability of the primers to amplify each identified exon at acceptable levels and the presence of primer sequences. These data were used to confirm that the performance of the pre-amp and PCR primers was adequate and consistent with previous primer lots.

This real-time reagent stability study established a 12-month shelf life of pre-amp and PCR primer mixes when stored at -20± 5°C. These data were confirmed by assessing the equivalence of MRD frequency in 40 clinical samples amplified with primer lots of different ages, and by tracking the stability of MRD measurements of synthetic molecules over time. The conditions tested in the real time stability study met the pre-specified of acceptance criteria of pairwise equivalence test of clinical specimens to be within 30% MRD frequency.

*j. Specimen Stability*

i. Frozen Bone Marrow Stability at -15° C to -25° C

To demonstrate frozen bone marrow stability, four bone marrow samples from donors were aliquoted and stored frozen (-15°C to -25°C). Samples were tested after freezing and specific values to clones within the specimen were compared to baseline. Frozen bone marrow is stable at -15°C to -25°C for 12 months within the prespecified ±30% MRD frequency variation.

ii. Bone Marrow Stability at Room Temperature and Refrigerated

To demonstrate bone marrow stability, four bone marrow samples from donors were aliquoted and stored at room temperature (19°C to 25°C) or refrigerated (2°C to 8°C) for up to seven days. Bone marrow samples stored at room temperature (19°C to 25°C) remained stable for three days. For bone marrow samples stored refrigerated (2°C to 8°C), MRD measurements for the clonoSEQ Assay remained stable for seven days. Acceptance criteria were based on the prespecified ±30% MRD frequency variation.

iii. Shipping Stability

Sample stability of bone marrow samples stored in Adaptive shipping containers at ambient temperature was tested under summer and winter shipping conditions. Study results demonstrated that samples are stable for up to 4 days (96 hours) under ambient shipping conditions based on an allowable 30% MRD frequency variation.

iv. Freeze/Thaw Stability of Bone Marrow Samples

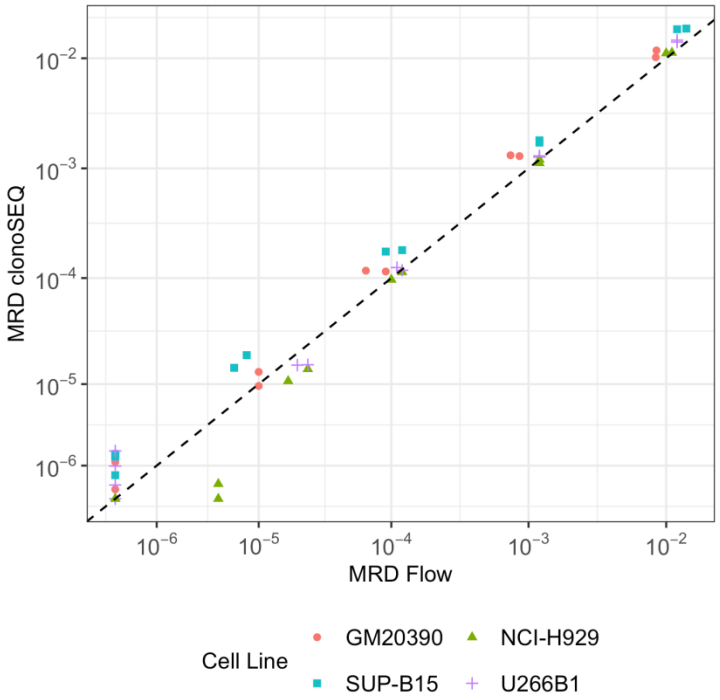
The stability of bone marrow samples was evaluated using four bone marrow samples from donors split into aliquots (0.25 ml) with one aliquot extracted upon receipt. The remaining aliquots were subjected to up 5 to freeze/thaw cycles. gDNA was extracted and the concentration was determined using a spectrophotometer. Each sample was processed using the clonoSEQ Assay. Bone marrow samples subjected to up to three freeze/thaw cycles continued to report acceptable sample MRD frequency within the prespecified ± 30% frequency variation.

2. Comparison Studies (Accuracy of Quantitative Measurement)

   *a. Assessment of clonoSEQ Assay Accuracy in Cell Mixtures Comparing to Multiparametric Flow Cytometry (mpFC)*

   The quantitative accuracy of the clonoSEQ Assay was compared to measurements obtained with a multiparametric flow cytometry (mpFC) method. Flow cytometry is considered the reference standard for MRD. Results were assessed using cell line blends. Measured MRD frequencies were compared against known frequencies based on diluting cell lines into background mononuclear cells at specific MRD levels. This study evaluated 2 MM cell lines (U266B1 and NCI-H929) and 2 ALL cell lines (SUP-B15 and GM20390). Each cell line was tested at 5 dilutions, from $5x10^{-7}$ to $1x10^{-2}$. Two replicates of each sample were assessed by the clonoSEQ assay and a validated mpFC assay. A pairwise comparison of MRD frequency measurements is shown in Figure 6. This study demonstrated similar quantitative accuracy when comparing clonoSEQ with mpFC at frequencies above $1x10^{-4}$.

**Figure 6: Pairwise comparison of MRD frequency measurements with mpFC and the clonoSEQ Assay**




*b. Concordance with mpFC in clinical samples*

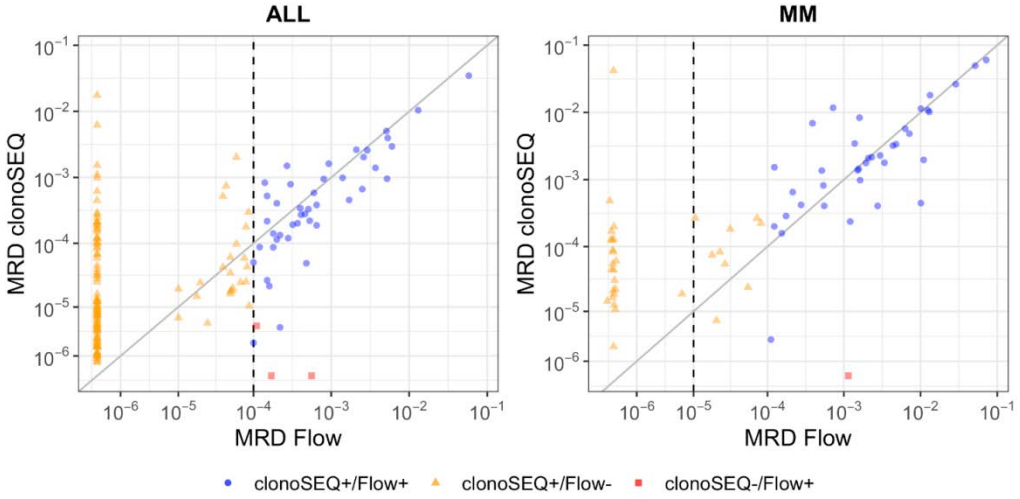
Two concordance studies between mpFC and the clonoSEQ Assay were performed using clinical samples. For both studies, concordance was assessed two ways: concordance of MRD positive or negative calls and concordance of reported MRD frequency. One study used 273 ALL samples from the Children's Oncology Group (COG) AALL0331 (standard risk) and AALL0232 (high risk) regimens and compared the clonoSEQ Assay to a validated mpFC assay. The other study performed a similar comparison using 91 MM samples from the Dana-Farber Cancer Institute (DFCI) Study 10-106 that were measured by both the clonoSEQ Assay and mpFC. MRD negativity was defined as $< 1x10^{-4}$ for mpFC in ALL (a commonly used threshold in that patient population) and $<1x10^{-5}$ for mpFC in MM. For the clonoSEQ Assay, MRD calls were assessed at the LOD in both studies. The negative percent agreement (NPA) between the clonoSEQ Assay and mpFC was 93.5% for ALL and 97.9% for MM (Table 10).

**Table 10: Summary of mpFC vs. the clonoSEQ Assay Concordance Data for ALL and MM**

| | clonoSEQ+ mpFC+ | clonoSEQ+ mpFC- | clonoSEQ- mpFC+ | clonoSEQ- mpFC- | mpFC reference | |
|---|---|---|---|---|---|---|
| | | | | | PPA | NPA |
| ALL | 43 | 117 | 3 | 110 | 48.5% (41.8-55.2%) | 93.5% (82.1-98.6%) |
| MM | 46 | 23 | 1 | 21 | 47.7% (32.5-63.3%) | 97.9% (88.7-99.9%) |

Concordance of MRD frequency was visualized by plotting reported MRD frequency of mpFC against the clonoSEQ Assay for both MM and ALL (Figure 7). Concordance of MRD call is indicated by color/shape; blue circles indicate samples had concordant MRD positive calls, while orange triangle and red squares denote discordant calls, with orange triangles indicating that clonoSEQ identified the sample as MRD positive and red squares indicating that mpFC identified the sample as MRD positive. To simplify the plot, samples with concordant MRD negative calls were not plotted. To quantify the similarity of reported MRD frequencies, correlations were calculated for samples with either concordant calls or mpFC positive calls; MRD frequencies were highly concordant (ALL, concordance correlation coefficient = 92.8%; MM, concordance correlation coefficient = 91.9%). These data demonstrate that at high disease burdens mpFC and clonoSEQ report similar MRD levels, while clonoSEQ continues to detect MRD at lower frequencies.

**Figure 7: Measurements of the clonoSEQ Assay Compared to Flow Cytometry Measurements from ALL (left) and MM (right) Clinical Studies.**

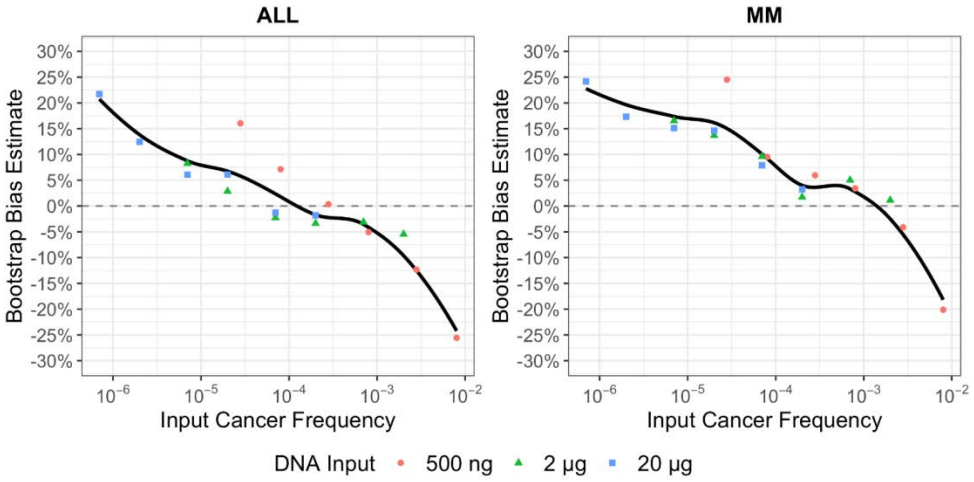


*c. Analysis of Quantitation Bias on Clinical Specimens*

The precision study described in section (L)(1)(b) evaluated blended gDNA extracted from 23 MM and 21 ALL specimens at 3 DNA inputs and 6 MRD frequencies per DNA input. These data were reanalyzed to evaluate if there was a quantitation bias for clonoSEQ. Sample MRD frequencies measured with the clonoSEQ Assay were compared to the expected MRD, value as calculated using flow cytometry on the original clinical sample and applying the appropriate dilution factor.

For ALL and MM, the quantitative accuracy of the clonoSEQ Assay was within ±25% across all tested diseased cell inputs (Figure 8). The assay tended to have a modest upward bias in MRD estimation at lower MRD frequencies and a modest downward bias at higher MRD frequencies.

**Figure 8: Estimates of Bias the clonoSEQ assay by MRD Frequency using clinical specimens**




3. Clinical Performance

Clinical validation for the measurement of MRD in ALL and MM was demonstrated using an analysis of samples obtained from two clinical studies for each disease indication.

*a. Clinical Validation of the clonoSEQ Assay for Acute Lymphoblastic Leukemia in Children's Oncology Group (COG) Studies AALL0232 and AALL0331*

The primary objective of this analysis was to establish the ability of the clonoSEQ Assay to predict event-free survival (EFS) at the MRD threshold of $10^{-4}$ using available bone marrow samples from patients who were enrolled in previously conducted COG studies AALL0232 and AALL0331.

COG study AALL0331 is a Phase III randomized study of different combination

chemotherapy regimens in pediatric patients with newly diagnosed standard risk B-precursor acute lymphoblastic leukemia. COG study AALL0232 is a Phase III randomized study of dexamethasone versus prednisone during induction and high-dose methotrexate with leucovorin rescue versus escalating-dose methotrexate without leucovorin rescue during interim maintenance I in patients with newly diagnosed high-risk acute lymphoblastic leukemia. Within these studies, bone marrow was collected at six separate time points to assess response to treatment, however, only the post induction marrow was used for MRD determination for these analyses.

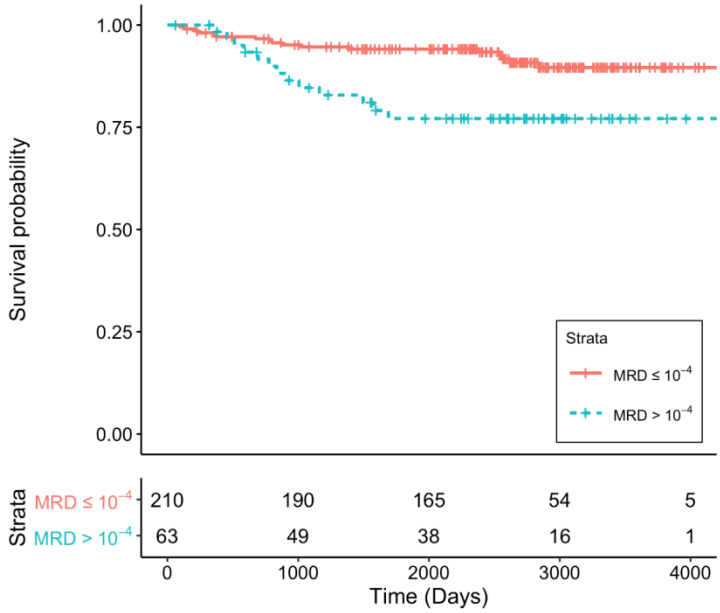
Clinical samples (pre-treatment BMA and day 29 post-induction BMA) were collected from 619 individuals, with samples from 315 patients that were enrolled as part of the "high risk" COG protocol AALL0232 and samples from 304 patients enrolled as part of the "standard risk" COG protocol AALL0331. Available specimens from these trials were tested with the clonoSEQ assay and results from both studies were pooled into a single analysis. Specimens were selected based on having a sufficient quantity of gDNA, available MRD flow cytometry results, and patients with study related endpoints for EFS.

A subset of 283 of the 619 patients originally enrolled in COG studies AALL0232 and AALL0331 had leftover samples of sufficient amount that could be tested with the clonoSEQ Assay. The population characteristics between these 283 patients were compared against the remaining 336 that were not tested and there were no significant differences in any characteristic that was evaluated, including age, gender, presence of specific genetic fusions, trisomy, and progression free survival. The 283 specimens were tested to evaluate the clinical performance of the clonoSEQ Assay and to demonstrate concordance in MRD measurements between the clonoSEQ Assay and results of original testing with a previous version of the clonoSEQ Assay and mpFC. Ten specimens did not pass QC, leaving results from 273 specimens available for the final analysis.

The clonoSEQ Assay MRD negativity of $< 1x10^{-4}$ was found to predict improved event free survival (EFS) irrespective of age (P=0.0034; Figure 9). Results demonstrate a 2.74-fold higher event risk in MRD positive patients ($MRD>10^{-4}$) compared to MRD negative patients (95% CI: 1.330-5.656). Similar findings were published in a broader COG analysis of the relationship between EFS and MRD negativity by an earlier version of the clonoSEQ Assay in pediatric ALL (Wood et al. 2018).[1]

[1] Wood B, Crossley B , Dai Y, et.al., Measurable residual disease detection by high-throughput sequencing improves risk stratification for pediatric B-ALL. Blood 2018; 131:1350-1359.

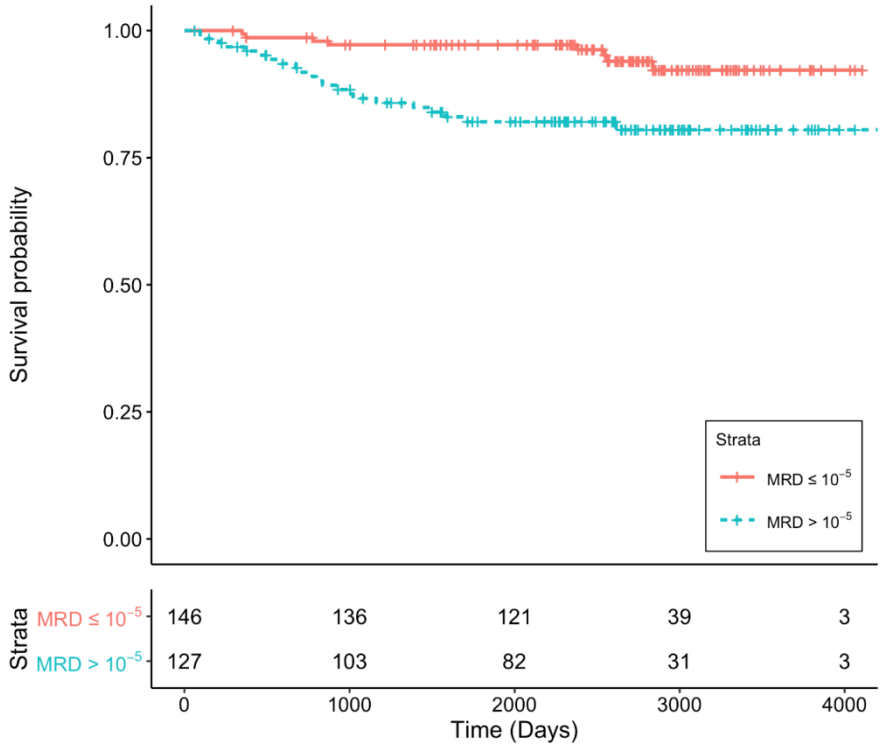
**Figure 9: Kaplan-Meier Survival Curve for EFS using the clonoSEQ Assay at an MRD cutoff of $1x10^{-4}$ in ALL**




The study was also designed to evaluate the clinical validity of the clonoSEQ Assay using alternative MRD thresholds and continuous MRD measures. Cox regression analysis of MRD and EFS using continuous MRD values demonstrates that the clonoSEQ Assay is significantly associated with EFS after adjusting for age (P=0.0057), and that each 10-fold increase in MRD level is associated with a 1.499-fold increase in event risk (95% CI: 1.139-1.974). These data further demonstrate that the MRD level remains a significant predictor of EFS, even after accounting for age, gender, and genetic abnormalities, which demonstrates the validity of MRD measurement in ALL.

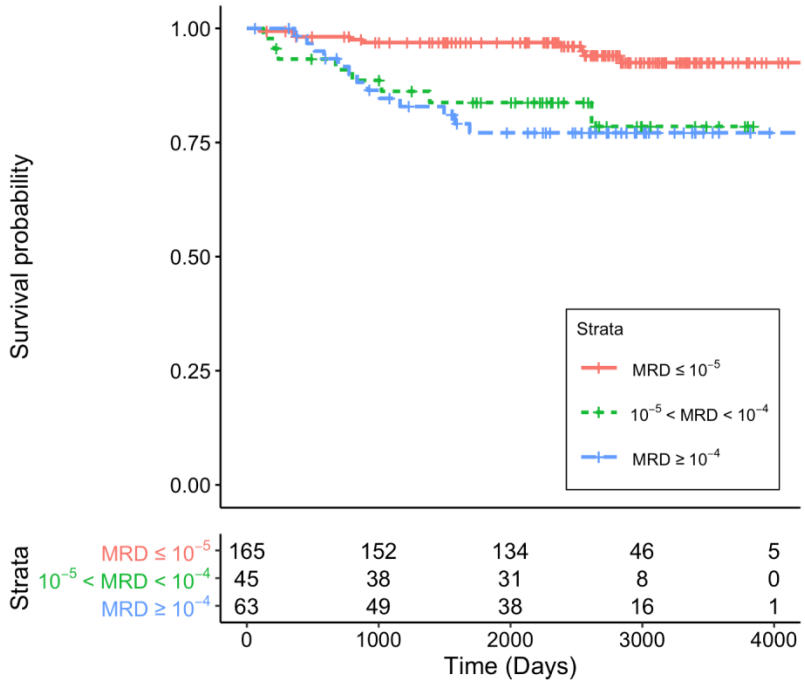
A qualitative assessment of MRD was also evaluated with MRD negativity defined as $< 1X10^{-5}$ (Figure 10). This threshold is significantly associated with EFS ($P=8.4x10^{-4}$).

**Figure 10: Kaplan-Meier Survival Curve for EFS using the clonoSEQ Assay at an MRD cutoff of $1x10^{-5}$ in ALL**




The clonoSEQ Assay was used to assess MRD at various disease burden thresholds to determine the correlation of MRD level with EFS. Patients who are clonoSEQ MRD negative ($<1x10^{-5}$) have longer EFS, followed by patients with MRD between $10^{-5}$ – $10^{-4}$ and, patients with MRD $>10^{-4}$ (P=0.00065; Figure 11). These data demonstrate that patients with the lowest levels of MRD have better outcomes than patients with higher disease burden, regardless of risk stratification.

**Figure 11: Non-parametric Kaplan-Meier survival curve for EFS in clonoSEQ MRD positive and MRD negative patients using three MRD bins: $\leq 10^{-5}$, $10^{-5} - 10^{-4}$, $\geq 10^{-4}$.**




These analyses demonstrated that MRD estimation by the clonoSEQ assay is associated with patient outcomes for B-cell precursor ALL.

*b. Clinical Validation of the clonoSEQ Assay for Multiple Myeloma*

Two separate studies were analyzed to support that MRD as estimated with the clonoSEQ Assay is associated with patient outcomes in MM. Samples for the analysis of the clonoSEQ Assay performance in MM were obtained from an ongoing randomized, open label, Phase III study of lenalidomide and bortezomib in a combination therapy regimen (DFCI Study 10-106). Multiple timepoints were assessed in this two-arm analysis and not all patients have the same number of MRD assessments (see data in section L.3.b). Patients on Arm A (blinded to Adaptive Biotechnologies) had assessments after eight cycles of RVD (lenalidomide, bortezomib, and dexamethasone), and then after lenalidomide maintenance. Patients on Arm B (blinded to Adaptive Biotechnologies) were assessed following 3 cycles of RVD, following auto transplant, and again after two more cycles of RVD consolidation, and then following lenalidomide maintenance.

i. <u>DFCI Study 10-106</u>

The objective of this study was to establish that the clonoSEQ Assay is predictive

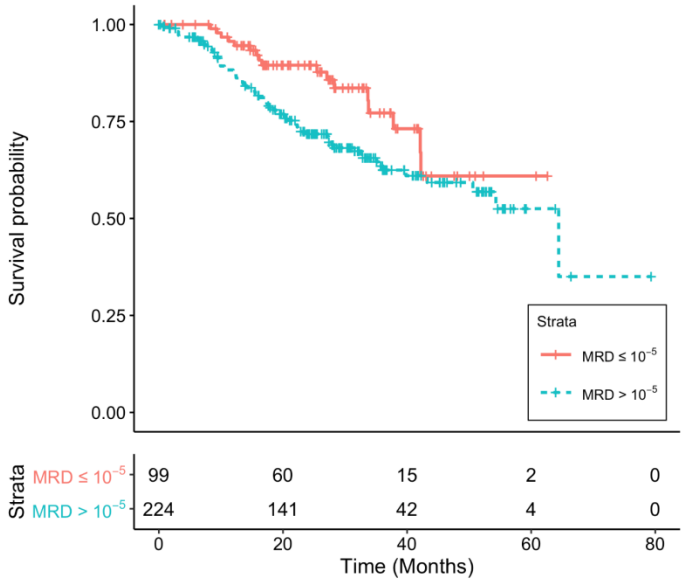
of progression-free survival (PFS) and disease-free survival (DFS) in MM. Patient samples were accrued under DFCI Study 10-106, "A Randomized Phase III Study Comparing Conventional Dose Treatment Using a Combination of Lenalidomide, Bortezomib, and Dexamethasone (RVD) to High-Dose Treatment with Peripheral Stem Cell Transplant in the Initial Management of Myeloma in Patients up to 65 Years of Age."

A subset of 365 of the 720 patients originally enrolled in DFCI Study 10-106 had leftover samples of sufficient amount to be tested with the ClonoSEQ Assay. The populations characteristics between these 365 patients were compared against the remaining 355 patients that were not tested and there were no significant differences in any characteristic that was evaluated, including age, gender, ISS staging, cytogenetic status and progression free survival. Samples from 365 patients were tested and results from 323 patients were evaluable and passed QC. Seventy-five (75) of these samples were from patients in complete response (CR) at the time of first MRD assessment. This study aimed to demonstrate the association of the first MRD measurement with PFS in patients who achieved CR and with PFS in all evaluable patients.

Samples from 75 patients who had achieved CR were evaluable for analysis. Continuous clonoSEQ MRD levels were modestly associated with DFS in patients who have achieved CR (P=0.064), such that patient with lower MRD levels were less likely to progress.

The ability of the clonoSEQ Assay MRD measurements to predict PFS in all 323 evaluable patients was also assessed. clonoSEQ measurements demonstrated that MRD status at a threshold of $1x10^{-5}$ at the time of first MRD measurement significantly predicts PFS in all patients (P= 0.027, Figure 12).

**Figure 12: Kaplan-Meier Survival Curve for PFS using the clonoSEQ Assay at an MRD cutoff of $10^{-5}$ in MM**




Cox regression analysis using a continuous measure of MRD was also associated with disease progression ($P=1.9x10^{-7}$). For every 10-fold increase in continuous clonoSEQ MRD measurement, the likelihood of an event is 1.69 times higher (95% CI:1.071-2.67).

ii. <u>ALCYONE study</u>

The ALCYONE Trial was a multicenter, randomized, open-label, active-controlled phase 3 trial that evaluated daratumumab plus bortezomib, melphan and prednisone (D-VMP) versus bortezomib, melphan and prednisone (VMP) in 706 patients with newly diagnosed multiple myeloma who were ineligible for stem-cell transplantation. The result of this study was reported in Mateos et al. 2018.[2]

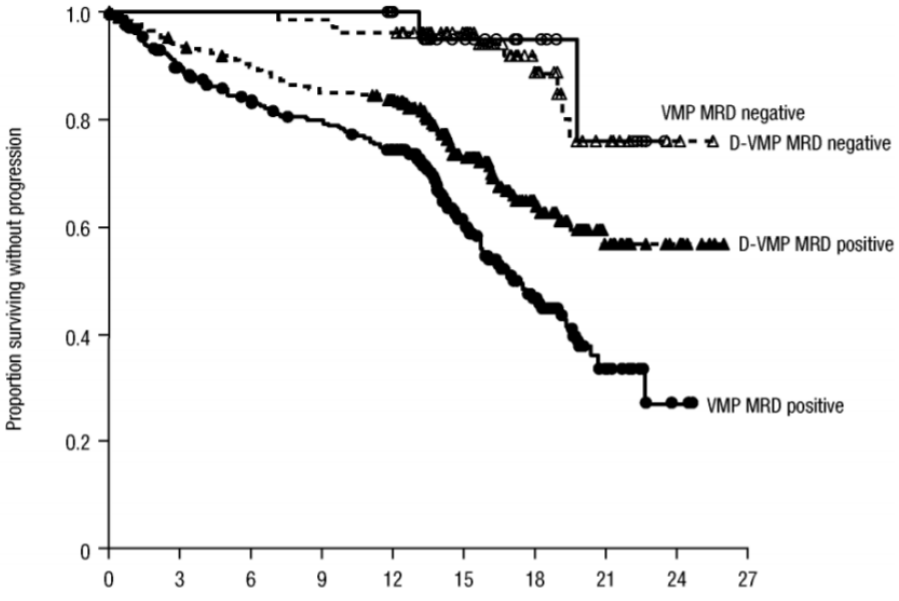
Within this trial, minimal residual disease was assessed by means of the clonoSEQ assay using bone marrow aspirate collected at screening, at the time of confirmation of complete response or stringent complete response, and at 12, 18, 24, and 30 months after the first dose in patients having a complete response or stringent complete response. Patients who did not achieve a CR were considered to be MRD positive. An MRD threshold of $10^{-5}$ was used for analysis.

Regardless of treatment group, patients who were MRD negative by the clonoSEQ Assay at $\leq 10^{-5}$ had longer progression-free survival (PFS) compared to

[2] Mateos MV, Dimopoulos MA, Cavo M, et al., Daratumumab plus Bortezomib, Melphalan, and Prednisone for Untreated Myeloma. N Engl J Med 2018; 378:518-528.

MRD positive patients (Figure 13).

**Figure 13: Analysis of MRD with Progression-Free Survival. Patients who were MRD negative by the clonoSEQ Assay had longer PFS compared to MRD positive patients**.




| Patients at risk | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| VMP MRD negative | 22 | 22 | 22 | 22 | 21 | 14 | 8 | 4 | 0 | 0 |
| D-VMP MRD negative | 78 | 78 | 78 | 77 | 75 | 58 | 31 | 14 | 2 | 0 |
| VMP MRD positive | 334 | 281 | 254 | 239 | 210 | 113 | 53 | 14 | 2 | 0 |
| D-VMP MRD positive | 272 | 244 | 234 | 221 | 210 | 121 | 62 | 21 | 8 | 0 |

*c. Clinical Cutoff:*

Not Applicable

## M. Instrument Name

Illumina NextSeq™ 500 System (qualified by Adaptive Biotechnologies, Inc.)

## N. System Descriptions

### 1. Modes of operation

The Illumina NextSeq™ 500 System is a high throughput sequencing system using Sequencing-By-Synthesis chemistry.

### 2. Software

The Adaptive clonoSEQ assay utilizes multiple software applications for performing and

analyzing sequencing results. The principle software applications used with the instrument include Illumina's (b) Fastq (b) (4), NextSeq Software System (b) (4), and the Adaptive clonoSEQ bioinformatics pipeline (b) (4)

3. **Level of concern**
   Moderate

4. **Specimen Handling**
   Refer to Device Description section above.

5. **Calibration and Quality Controls**
   Refer to Device Description section above.

## O. Other Supportive Instrument Performance Characteristics Data Not Covered In The "Performance Characteristics" Section Above:

The labeling supports the decision to grant the De Novo request for this device.

## P. Patient Perspective

This submission did not include specific information on patient perspectives for this device.

## Q. Identified Risks to Health and Identified Mitigations:

| Identified Risks to Health | Identified Mitigations |
|---|---|
| **Incorrect test results** | General controls and special controls (1), (2) and (3) |
| **Incorrect interpretation of test results** | General controls and special controls (1), (2), (3) and (4) |

## R. Benefit Risk Summary

**Summary of Assessment of Benefit**
The Adaptive Biotechnologies clonoSEQ Assay is a Next generation sequencing-based in vitro diagnostic to evaluate minimal residual disease, evaluating rearranged IgH (VDJ), IgH(DJ), IgK, and IgL receptor gene sequences, as well as translocated BCL1/IgH(J) and BCL2/IgH(J) sequences. This assay measures minimal residual disease (MRD) to monitor changes in burden of disease during and after treatment. There is significant potential benefit associated with the use of this device in the clinical setting to assist clinicians manage patients with ALL or MM in accordance with professional guidelines and in conjunction with other clinicopathological features.

**Summary of the Assessment of Risk**
**For the Proposed Indications for Use:**
There is potential risk associated with the use of this device, mainly due to 1) false positives, false negatives, and failure to provide a result and 2) incorrect interpretation of test results by the user.

The probable risk associated with the use of this device is that false positives, false negatives, and failure to provide a result or incorrect interpretation of test results by the user can result in mismanagement of patients, resulting in significant consequences. Mismanagement of patients can occur because the marker measurement from the device is erroneous or interpreted erroneously, diverting patients away from the correct management as per professional guidelines. The risk of mismanagement is partially attenuated because the results of this device are to be use with other clinicopathological factors. In addition, the analytical validation of this device, in part mitigates the risks associated with this device.

There is moderate uncertainty about risk due to insufficient experience with the use of this device.

**Summary of the Assessment of Benefit-Risk, considering risk mitigation strategies**
**For the Proposed Indications for Use:**
In summary, the probable benefits of this device outweigh the probable risks associated with the device, when considering the mitigations provided by the limitations and special controls, beyond general controls.

**S. Conclusion:**

The information provided in this de novo submission is sufficient to classify this device into class II under regulation 21 CFR 866.6100. FDA believes that the class II device's stated special controls, in combination with the applicable general controls, provide a reasonable assurance of the safety and effectiveness of the device type. The device is classified under the following:

| | |
|---|---|
| Product Code: | QDC |
| Device Type: | DNA-based test to measure minimal residual disease in hematological malignancies |
| Class: | II (Special Controls) |
| Regulation: | 21 CFR 866.6100 |

(a) Identification: A DNA-based test to measure minimal residual disease in hematological malignancies is a prescription in vitro diagnostic device that identifies and quantifies specific nucleic acid sequences within human tissues to estimate the percentage of cells that harbor the specific sequence(s). The test is intended to be used as an aid to measure minimal residual disease to assess the change in burden of disease during monitoring of

treatment. The test is indicated for use by qualified healthcare professionals in accordance with professional guidelines for clinical decision-making, in conjunction with other clinicopathological features.

(b) Classification: Class II (special controls). A DNA-based test to measure minimal residual disease in hematological malignancies must comply with the following special controls:

(1) Design verification and validation must include:

(i) A detailed description of the device including:

(A) A detailed description of all test components, reagents, instrumentation, and software, including, but not limited to, software applications and any hardware-based devices that incorporate software.

(B) A detailed description of all genomic regions that are detected and quantified by the assay.

(C) A detailed description of the methodology and protocols for each step of the test, including description of the quality metrics, thresholds, and filters at each step of the test that are implemented for final result reporting and a description of the metrics for run-failures, specimen-failures, and invalids, as appropriate.

(D) Detailed specifications and procedures for sample collection, processing, and storage.

(E) A description of the internal and external controls that are recommended or provided. The description must identify those control elements that are incorporated into the testing procedure. If appropriate, this description must include a description of the controls and control procedures used during the sequencing and data analysis.

(ii) Identification of risk mitigation elements used by the device, including a detailed description of all additional procedures, methods, and practices incorporated into the instructions for use that mitigate risks associated with use of the device.

(iii) As part of the risk management activities, an appropriate end user device training program must be offered as an effort to mitigate the risk of failure from user error, as appropriate.

(iv) Description of analytical and clinical studies including:

(A) Device performance data that demonstrates the ability to measure minimal residual disease in the claimed specimen type(s) from patients that are representative of the intended use population. Data can be obtained via:

(*1*) A method comparison study comparing the device to a predicate device with clinical data for the specified hematological neoplastic indication using the specified specimen type(s), or
(*2*) A clinical study demonstrating clinical validity using well-characterized clinical specimens from patients with known clinical outcomes using a study design deemed acceptable by the FDA.

(B) Device precision (repeatability and reproducibility) data using clinical samples covering the range of minimal residual disease frequencies reported by the test and covering the stated range of DNA inputs that are indicated as allowable for use with the test. Results shall be reported as the standard deviation and/or percentage coefficient of variation with the 95% confidence interval for each level tested. The study must evaluate all sources of variability including, as appropriate, between-site and between operator (minimum of 3 sites of which 2 must be external with a minimum of 2 operators per site), between-day (minimum of 3 days), between-run, within-run, between-lot (minimum of 3 lots), between instrument (minimum of 3 instruments), and total variation.

(C) Device linearity data generated from samples covering the device measuring range using a dilution panel created from clinical samples.

(D) Device accuracy by comparison to flow cytometry across the measuring interval or to the predicate method across the measuring interval.

(E) Device analytic sensitivity data, including limit of blank, limit of detection, and limit of quantitation, using a dilution panel created from clinical samples.

(F) Analytical specificity data, including interference and cross-contamination, and index cross-contamination, as appropriate.

(G) Validation of pre-analytical methods, including DNA extraction methods and cell enrichment methods, as appropriate.

(H) Device stability data, including real-time stability of reagents under various storage times and temperatures.

(I) Specimen and prepared sample stability data established for each specimen matrix in the anticoagulant combinations and storage/use conditions that will be indicated, including specimen transport, as appropriate.

(2) The intended use required on the label under 21 CFR 809.10(a)(4) and on the labeling under 21 CFR 809.10(b)(5)(ii), as applicable, must include:

(i) The clinical hematopoietic malignancy for which the assay was designed and validated (e.g., multiple myeloma or B-cell acute lymphoblastic leukemia);

(ii) Specimen type (e.g., bone morrow);

(iii) The specific DNA regions that are being identified and quantified (e.g., rearranged IgH (VDJ), IgH(DJ), IgK, and IgL receptor gene sequences); and

(iv) A statement that the results are indicated to be interpreted by qualified healthcare professionals in accordance with professional guidelines for clinical decision-making in conjunction with other clinicopathological features.

(3) The 21 CFR 809.10(b) labeling must include information that demonstrates the performance characteristics of the test, including a detailed summary of the performance studies conducted and their results, as described in (b)(1)(iv)(A) through (b)(1)(iv)(I).

(4) The device output, including any test report, must include the estimated minimal residual disease (MRD) frequency and an appropriate range of the uncertainty of that frequency based on the amount of DNA that was evaluated by the test and the number of specific nucleic acid sequences that were detected (e.g., "MRD = $1.2 \times 10^{-5}$ [Range = $0.8 \times 10^{-6}$ to $2.0 \times 10^{-5}$]").