**SDC, Figure 1**. Gating strategy to evaluate CD154+T-cytotoxic memory cells (CD154+TcM) by flow cytometry in each reaction condition. SDC, 1a. Scatterplot shows the location of lymphocytes on the basis of forward and side scatter. SDC, 1b. T-cells (CD3+) are gated from total lymphocytes using side scatter by CD3 in the FITC channel. SDC, 1c. Live T-cells are identified after exclusion of dead cells which take up the dye 7-AAD (CD3+7AAD+). SDC, Panel 1d. By gating further on live CD3 cells, responder T-cytotoxic cells (Tc or CD8) which have been pre-labeled with anti-CD8-APCH7 are separated from stimulator Tc which have been pre-labeled with anti-CD8-PE-Cy7, and from T-cells which do not express the CD8 marker (Q3). SDC, 1e. By gating further on responder Tc, those with the memory marker CD45RO (TcM) and are labeled with APC-labeled anti-CD45RO are separated from those that do not express this marker (naïve). SDC, 1f-1h. By gating further on responder TcM, those that express CD154 and (CD154+TcM) and are labeled with PE-labeled anti-CD154 are separated from those that do not express this marker. SDC, 1f. The area occupied by CD154+TcM is identified by a gate which excludes all cells in the negative control reaction. In this reaction, recipient cells are incubated alone without fluorescent antibody to CD154. This scatterplot configuration is fixed and used to analyze all other reaction conditions. SDC, 1g. The background reaction in which recipient cells are incubated with anti-CD154-PE without stimulator shows “background” counts of CD154+TcM. SDC, 1h. The addition of HLA-non-identical stimulator cells in this stimulated reaction results in stimulation-induced increase in CD154+TcM counts.

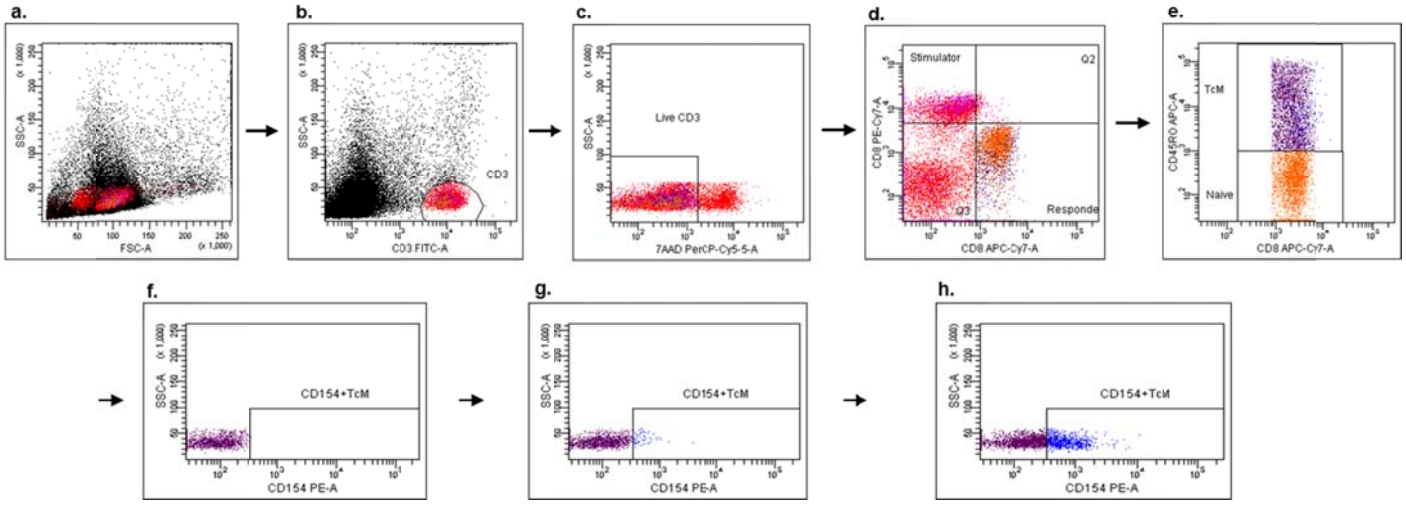
FSC: Forward scatter, SSC: Side scatter, A: Area, FITC: Fluorescein isothiocyanate , PerCP-Cy5.5: Peridinin chlorophyll protein and cyanine 5.5 tandem dye, APCH7: allophycocyanin and cyanine

**Supplemental Digital Content (SDC)**

**Figure**

7 tandem dye, PECy7: phytoerythrin and cyanine 7 tandem dye, APC: allophycocyanin, PE: Phycoerythrin.

SDC, Figure 1



**Supplemental Digital Content (SDC)**

**Figure**

**SDC Figure 2**. General approach to training-set/validation-set testing maximizes utilization of samples from this rare test population. A subject e.g. subjects X or Y, can provide up to one pre-transplant and one post-transplant sample in either training set or the validation set. Arrows indicate that a sample used for model building in the training set can only be used to predict outcomes of independent validation set samples for the corresponding time period. Therefore, if a subject, e.g. Subject Z, provided samples for model building in the training set, subsequent samples from a different time period could still be included in the validation set. Strikethroughs imply unacceptable samples.

# Figure S2

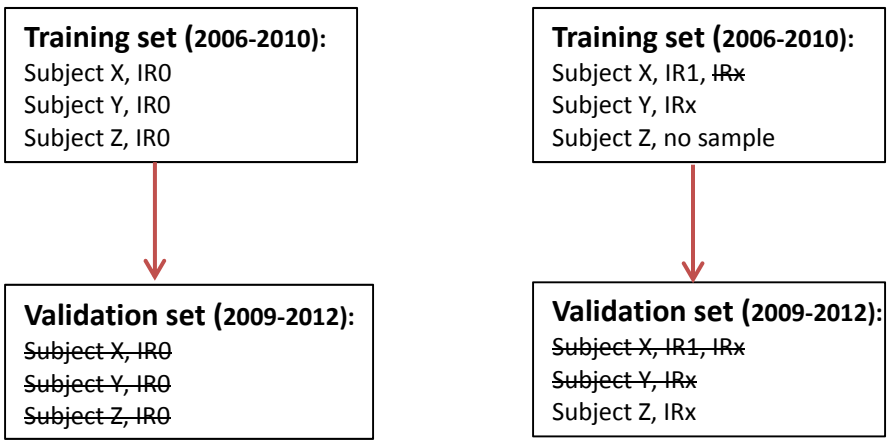


**Table S1**: Mean immunoreactivity index (IR) for allospecific CD154+T-cytotoxic memory cells (CD154+TcM) in training and validation set samples R=Rejector, NR=Non-rejector, IR0=pre-transplant samples, IR1+IRx=post-transplant samples obtained within the first 60 days after transplantation (IR1), and days 200-onward after transplantation (IRx), CI=confidence interval, low=lower, up=upper, SD=standard deviation, Min=minimum, Max=maximum.

| Data Type | Sample type | Outcome | N | Mean | CI-low | CI-up | SD | Median | Min | Max |
|---|---|---|---|---|---|---|---|---|---|---|
| Training | IR0 | R + NR | 49 | 3.1 | 1.1 | 5.2 | 7.1 | 1.4 | 0.06 | 46.4 |
| | | NR | 24 | 1.1 | 0.5 | 1.7 | 1.4 | 0.7 | 0.06 | 7.0 |
| | | R | 25 | 5.0 | 1.1 | 9.0 | 9.6 | 1.9 | 0.5 | 46.4 |
| | IR1 + IRx | R + NR | 98 | 1.6 | 1.0 | 2.3 | 3.2 | 0.7 | 0.03 | 26.4 |
| | | NR | 74 | 0.8 | 0.5 | 1.1 | 1.2 | 0.6 | 0.03 | 6.7 |
| | | R | 24 | 4.1 | 1.8 | 6.4 | 5.5 | 2.5 | 0.05 | 26.4 |
| | | | | | | | | | | |
| **Data Type** | **IR Level** | **Outcome** | **N** | **Mean** | **CI-low** | **CI-up** | **SD** | **Median** | **Min** | **Max** |
| Validation | IR0 | R + NR | 33 | 1.2 | 0.7 | 1.6 | 1.2 | 0.9 | 0.01 | 7.1 |
| | | NR | 19 | 0.7 | 0.6 | 0.9 | 0.3 | 0.8 | 0.01 | 1.5 |
| | | R | 14 | 1.7 | 0.8 | 2.7 | 1.6 | 1.4 | 0.4 | 7.1 |
| | IR1 + IRx | R + NR | 64 | 1.2 | 1.0 | 1.4 | 0.8 | 0.9 | 0.3 | 3.9 |
| | | NR | 45 | 1.0 | 0.8 | 1.3 | 0.8 | 0.8 | 0.3 | 3.9 |
| | | R | 19 | 1.6 | 1.2 | 2.0 | 0.8 | 1.4 | 0.5 | 3.7 |

**Supplemental Digital Content**
**Tables**

The following additional studies were performed to determine test variability.

1. To document variability if three-different operators and three different flow cytometers tested the same sample on the same day, assays between twenty one unique pairs of HLA-mismatched PBL from normal human subjects were performed by three different operators and instruments. Variability in CD154+TcM generated in the stimulated reaction was determined. Results are summarized in SDC, Tables 2a and 2b.
2. To determine variability in test results between two technologists, five samples were assayed simultaneously by two different technologists on the same day. Variability in CD154+TcM generated in the stimulated reaction was determined. Results are summarized in SDC, Tables 3a and 3b.

**Supplemental Digital Content**
**Tables**

**Table S2a:** Mean %CD154TcM on same-day testing by three different technologists/instruments (1a, 1b, 1c)

| Run | Reaction | N | Mean | CI-low | CI-up | SD | Median | Min | Max |
|---|---|---|---|---|---|---|---|---|---|
| 1a | Background | 21 | 4.0 | 2.3 | 5.7 | 3.8 | 2.9 | 0.5 | 13.2 |
| | Stimulated | 21 | 24.5 | 19.5 | 29.6 | 11.0 | 24.7 | 9.0 | 55.0 |
| 1b | Background | 21 | 5.0 | 2.8 | 7.2 | 4.9 | 3.5 | 0.9 | 19.2 |
| | Stimulated | 21 | 22.8 | 18.0 | 27.6 | 10.5 | 22.4 | 9.8 | 52.2 |
| 1c | Background | 21 | 5.8 | 3.2 | 8.3 | 5.7 | 3.4 | 1.2 | 18.4 |
| | Stimulated | 21 | 22.9 | 18.0 | 27.8 | 10.8 | 22.4 | 9.8 | 51.7 |

**Table S2b.** Mean %CV for %CD154+TcM in samples assayed by three technologists on three instruments

| Run | Reaction | N | Mean | CI-low | CI-up | SD | Median | Min | Max |
|---|---|---|---|---|---|---|---|---|---|
| 1a, 1b & 1c | Background | 21 | 34.3 | 26.1 | 42.2 | 17.9 | 32.2 | 7.7 | 76.3 |
| | Stimulated | 21 | 8.2 | 6.0 | 10.4 | 4.8 | 6.5 | 2.4 | 20.7 |
| 1a & 1d | Background | 20 | 45.3 | 33.8 | 56.8 | 24.5 | 42.6 | 5.0 | 87.2 |
| | Stimulated | 20 | 8.9 | 5.6 | 12.2 | 7.0 | 6.5 | 0.8 | 24.9 |

**Supplemental Digital Content**
**Tables**

**Table S3a:** Mean %CD154+TcM for all (n=5) samples for each of two technicians

| Tech | Reaction | N | Mean | CI-low | CI-up | SD | Median | Min | Max |
|---|---|---|---|---|---|---|---|---|---|
| Tech 1 | Background | 5 | 1.2 | 0.6 | 1.9 | 0.5 | 1.3 | 0.4 | 1.7 |
| | Stimulated | 5 | 5.6 | 3.1 | 8.0 | 2.0 | 4.8 | 3.1 | 7.8 |
| Tech 2 | Background | 5 | 0.7 | 0 | 1.5 | 0.6 | 0.6 | 0.1 | 1.7 |
| | Stimulated | 5 | 5.8 | 3.2 | 8.4 | 2.1 | 5.3 | 3.0 | 8.2 |

**Table S3b:** Mean %CV for %CD154+TcM for all (n=5) samples for each of two technicians

| Tech | Reaction | N | Mean | CI-low | CI-up | SD | Median | Min | Max |
|---|---|---|---|---|---|---|---|---|---|
| Tech 1 and 2 | Background | 5 | 55.1 | 19.4 | 90.9 | 28.8 | 64.5 | 18.7 | 84.0 |
| | Stimulated | 5 | 4.8 | 1.1 | 8.5 | 3.0 | 4.2 | 1.8 | 8.4 |

**Supplemental Digital Content**
**Tables**

**Table S4**: Performance of optimal (multiple variable) post-transplant (7a upper table) and pre-transplant (7b lower table) models based on IR of CD154+TcM in training and validation sets.

| Cohort | AUC | Cut value | Sensitivity | 95% CI | Specificity | 95% CI | PPV | 95% CI | NPV | 95% CI |
|---|---|---|---|---|---|---|---|---|---|---|
| Training set (n=98) | 0.937 | -0.85 | **95%** (20/21) | 74%, 100% | **89%** (62/70) | 78%, 95% | **71%** (20/28) | 51%, 86% | **98%** (62/63) | 90%, 100% |
| Validation set (n=67) | 0.711 | -0.85 | **53%** (10/19) | 29%, 75% | **83%** (40/48) | 69%, 92% | **56%** (10/18) | 31%, 78% | **82%** (40/49) | 67%, 91% |

| Cohort | AUC | Cut value | Sensitivity | 95% CI | Specificity | 95% CI | PPV | 95% CI | NPV | 95% CI |
|---|---|---|---|---|---|---|---|---|---|---|
| Training set (n=49) | 0.913 | 0.121 | **88%** (22/25) | 68%, 97% | **88%** (21/24) | 67%, 97% | **88%** (22/25) | 68%, 97% | **88%** (21/24) | 67%, 97% |
| Validation set (n=33) | 0.761 | 0.121 | **57%** (8/14) | 30%, 81% | **79%** (15/19) | 54%, 93% | **67%** (8/12) | 35%, 89% | **71%** (15/21) | 48%, 88% |

**Supplemental Digital Content**
**Tables**

**Table S5:** Coefficients of variation for the multivariate (upper table) and single variable (lower table) models for post-transplant training set samples.

| IR1+IRx Training full model | Estimate | Std. Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|
| **(Intercept)** | -1.573 | 0.852 | -1.848 | 0.065 |
| **Race** | 1.561 | 0.812 | 1.923 | 0.054 |
| **IR** | 3.413 | 0.930 | 3.672 | 2.41E-04 |
| **Organ.1** | -1.397 | 0.491 | -2.846 | 0.004 |
| **Time between Transplant and sample days** | -0.001 | 0.000 | -1.861 | 0.063 |
| **FK level-day of sample** | 0.223 | 0.078 | 2.872 | 0.004 |

| IR1+IRx Training full model | Estimate | Std. Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|
| **(Intercept)** | -1.191 | 0.304 | -3.913 | 9.11E-05 |
| **IR** | 3.306 | 0.766 | 4.317 | 1.58E-05 |

**Supplemental Digital Content**
**Tables**

**Effect of confounders:** Comparable test performance within the range seen in training and validation set samples was also seen in samples sub-grouped by the time of sampling with respect to transplantation, type of organ transplanted, type of induction, whether actual or surrogate donor stimulators were used, and whether rejection or non-rejection were diagnosed by “for-cause” or surveillance biopsy or clinically (SDC, Tables 6-10 in Supplemental Digital Content: Tables). Performance estimates are less likely to be meaningful for those subgroups with small numbers of samples.

**Supplemental Digital Content**
**Tables**

**Table S6:** General test performance stratified by time after transplantation for training and validation set. R=rejector, NR=Non-rejector, AUC=Area under the receiver-operating characteristic curve, Sens=sensitivity, Spec=specificity, LCI=lower confidence interval, UCI=upper confidence interval, PPV=positive predictive value, NPV=negative predictive value.

| Dataset | Total (n) | R (n) | NR (n) | AUC | Sens n | 95% LCI | 95% UCI | Spec n | 95% LCI | 95% UCI | PPV n | 95% LCI | 95% UCI | NPV n | 95% LCI | 95% UCI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IR0 Validation | 33 | 14 | 19 | 0.842 | 57.1 8/14 | 29.6 | 81.2 | 89.5 17/19 | 65.5 | 98.2 | 80 8/10 | 44.2 | 96.5 | 73.9 17/23 | 51.3 | 88.9 |
| IR0 Training | 49 | 25 | 24 | 0.82 | 80 20/25 | 58.7 | 92.4 | 70.8 17/24 | 48.8 | 86.6 | 74.1 20/27 | 53.4 | 88.1 | 77.3 17/22 | 54.2 | 91.3 |
| IR1&IRx Training | 98 | 24 | 74 | 0.878 | 91.7 22/24 | 71.5 | 98.5 | 83.86 2/74 | 73 | 91 | 64.7 22/34 | 46.5 | 79.7 | 96.9 62/64 | 88.2 | 99.5 |
| IR1&IRx Validation | 64 | 19 | 45 | 0.791 | 84.2 16/19 | 59.5 | 95.8 | 8036/ 45 | 64.9 | 89.9 | 6416/ 25 | 42.6 | 81.3 | 92.3 36/39 | 78 | 98 |
| IR1 Training | 48 | 17 | 31 | 0.841 | 88.2 15/17 | 62.3 | 97.9 | 80.6 25/31 | 61.9 | 91.9 | 71.4 15/21 | 47.7 | 87.8 | 92.6 25/27 | 74.2 | 98.7 |
| IR1 Validation | 30 | 15 | 15 | 0.796 | 80 12/15 | 51.4 | 94.7 | 86.7 13/15 | 58.4 | 97.7 | 85.7 12/14 | 56.2 | 97.5 | 81.2 13/16 | 53.7 | 95 |
| IRx Training | 50 | 7 | 43 | 0.95 | 100 7/7 | 56.1 | 100 | 86 37/43 | 71.4 | 94.2 | 53.87/ 13 | 26.1 | 79.6 | 100 37/37 | 88.3 | 100 |
| IRx Validation | 34 | 4 | 30 | 0.858 | 100 4/4 | 39.6 | 100 | 76.7 23/30 | 57.3 | 89.4 | 36.44/ 11 | 12.4 | 68.4 | 100 23/23 | 82.2 | 100 |

**Supplemental Digital Content**
**Tables**

**Table S7: Test** performance stratified by type of donor stimulator, whether actual donor or surrogate donor. AUC=Area under the receiver-operating characteristic curve, Sens=sensitivity, Spec=specificity PPV=positive predictive value, NPV=negative predictive value.

| Dataset | Donor cell type | Total n | Rejector n | Non-rejector n | AUC | Sens | Spec | PPV | NPV |
|---|---|---|---|---|---|---|---|---|---|
| IR0 Training | Surrogate Donor | 40 | 20 | 20 | 0.87 | 90 | 70 | 75 | 87.5 |
| IR0 Validation | Surrogate Donor | 31 | 14 | 17 | 0.868 | 57.1 | 88.2 | 80 | 71.4 |
| IR1&IRx Training | Surrogate Donor | 80 | 19 | 61 | 0.847 | 89.5 | 80.3 | 58.6 | 96.1 |
| IR1&IRx Validation | Surrogate Donor | 59 | 17 | 42 | 0.817 | 88.2 | 81 | 65.2 | 94.4 |
| IR0 Training | Actual Donor | 9 | 5 | 4 | 0.75 | 40 | 75 | 66.7 | 50 |
| IR0 Validation | Actual Donor | 2 | 0 | 2 | NA | NA | 100 | NA | 100 |
| IR1&IRx Training | Actual Donor | 18 | 5 | 13 | 1 | 100 | 100 | 100 | 100 |
| IR1&IRx Validation | Actual Donor | 5 | 2 | 3 | 0.5 | 50 | 66.7 | 50 | 66.7 |

**Supplemental Digital Content**
**Tables**

**Table S8**: Test performance stratified by organ type. L=Liver, SB=small bowel or intestine, LSB=combined liver-intestine, LK=combined liver-kidney, LL=liver-lung. AUC=area under the receiver-operating –characteristic curve, sens=sensitivity, spec=specificity, PPV=positive predictive value, NPV=positive predictive value, NA=could not be calculated, n=number.

| Dataset | Organ | Total n | Rejector n | Non-rejector n | AUC | Sens | Spec | PPV | NPV |
|---|---|---|---|---|---|---|---|---|---|
| IR0 Training | L | 25 | 14 | 11 | 0.89 | 85.7 | 90.9 | 92.3 | 83.3 |
| IR0 Validation | L | 26 | 11 | 15 | 0.897 | 63.6 | 86.7 | 77.8 | 76.5 |
| IR1&IRx Training | L | 68 | 18 | 50 | 0.905 | 94.4 | 88 | 73.9 | 97.8 |
| IR1&IRx Validation | L | 54 | 17 | 37 | 0.755 | 82.4 | 75.7 | 60.9 | 90.3 |
| IR0 Training | LSB | 11 | 7 | 4 | 0.75 | 71.4 | 75 | 83.3 | 60 |
| IR0 Validation | LSB | 1 | 1 | 0 | NA | 0 | NA | NA | 0 |
| IR1&IRx Training | LSB | 15 | 3 | 12 | 0.833 | 66.7 | 75 | 40 | 90 |
| IR1&IRx Validation | LSB | 2 | 1 | 1 | 1 | 100 | 100 | 100 | 100 |
| IR0 Training | LL | 0 | NA | NA | NA | NA | NA | NA | NA |
| IR0 Validation | LL | 0 | NA | NA | NA | NA | NA | NA | NA |
| IR1&IRx Training | LL | 0 | NA | NA | NA | NA | NA | NA | NA |
| IR1&IRx Validation | LL | 1 | 0 | 1 | NA | NA | 100 | NA | 100 |
| IR0 Training | SB | 10 | 2 | 8 | 0.875 | 100 | 50 | 33.3 | 100 |
| IR0 Validation | SB | 6 | 2 | 4 | 1 | 50 | 100 | 100 | 80 |
| IR1&IRx Training | SB | 13 | 2 | 11 | 0.955 | 100 | 72.7 | 40 | 100 |
| IR1&IRx Validation | SB | 7 | 1 | 6 | 1 | 100 | 100 | 100 | 100 |
| IR0 Training | LK | 3 | 2 | 1 | 0 | 50 | 0 | 50 | 0 |
| IR0 Validation | LK | 0 | NA | NA | NA | NA | NA | NA | NA |
| IR1&IRx Training | LK | 2 | 1 | 1 | 1 | 100 | 100 | 100 | 100 |
| IR1&IRx Validation | LK | 0 | NA | NA | NA | NA | NA | NA | NA |

**Supplemental Digital Content**
**Tables**

**Table S9**: Test performance stratified by type of induction. None=no induction, thymo=thymoglobulin or rabbit anti-human thymocyte globulin, campath=alemtuzumab. Sens=sensitivity, Spec=specificity PPV=positive predictive value, NPV=negative predictive value

| Dataset | Induction Type | Total n | Rejector n | Non-rejector n | AUC | Sens | Spec | PPV | NPV |
|---|---|---|---|---|---|---|---|---|---|
| IR0 Training | None | 10 | 6 | 4 | 0.875 | 83.3 | 100 | 100 | 80 |
| IR0 Validation | None | 13 | 6 | 7 | 0.952 | 50 | 100 | 100 | 70 |
| IR1&IRx Training | None | 33 | 4 | 29 | 0.935 | 100 | 89.7 | 57.1 | 100 |
| IR1&IRx Validation | None | 29 | 10 | 19 | 0.787 | 80 | 84.2 | 72.7 | 88.9 |
| IR0 Training | Thymo | 29 | 15 | 14 | 0.786 | 73.3 | 64.3 | 68.8 | 69.2 |
| IR0 Validation | Thymo | 19 | 8 | 11 | 0.807 | 62.5 | 81.8 | 71.4 | 75 |
| IR1&IRx Training | Thymo | 54 | 17 | 37 | 0.859 | 88.2 | 81.1 | 68.2 | 93.8 |
| IR1&IRx Validation | Thymo | 32 | 9 | 23 | 0.78 | 88.9 | 73.9 | 57.1 | 94.4 |
| IR0 Training | Campath | 10 | 4 | 6 | 0.958 | 100 | 66.7 | 66.7 | 100 |
| IR0 Validation | Campath | 1 | 0 | 1 | NA | NA | 100 | NA | 100 |
| IR1&IRx Training | Campath | 11 | 3 | 8 | 0.917 | 100 | 75 | 60 | 100 |
| IR1&IRx Validation | Campath | 3 | 0 | 3 | NA | NA | 100 | NA | 100 |

**Supplemental Digital Content**
**Tables**

**Table S10**: Test performance stratified by whether outcome was determined without biopsy (no Bx), with for cause Bx or with surveillance Bx. Sens=sensitivity, Spec=specificity PPV=positive predictive value, NPV=negative predictive value

| Dataset | Bx type | Total n | Rejector n | Non-rejector n | AUC | Sens | Spec | PPV | NPV |
|---|---|---|---|---|---|---|---|---|---|
| IR0 Training | no Bx | 11 | 2 | 9 | 0.889 | 100 | 77.8 | 50 | 100 |
| IR0 Validation | no Bx | 12 | 1 | 11 | 1 | 100 | 81.8 | 33.3 | 100 |
| IR1&IRx Training | no Bx | 57 | **2** | 55 | 0.932 | 100 | 89.1 | 25 | 100 |
| IR1&IRx Validation | no Bx | 39 | 1 | 38 | 0.934 | 100 | 78.9 | 11.1 | 100 |
| IR0 Training | for cause Bx | 17 | 13 | 4 | 0.904 | 84.6 | 100 | 100 | 66.7 |
| IR0 Validation | for cause Bx | 11 | 10 | 1 | 0.7 | 60 | 100 | 100 | 20 |
| IR1&IRx Training | for cause Bx | 20 | 15 | 5 | 0.88 | 93.3 | 80 | 93.3 | 80 |
| IR1&IRx Validation | for cause Bx | 13 | 13 | 0 | NA | 76.9 | NA | 100 | NA |
| IR0 Training | Surveillance Bx | 21 | 10 | 11 | 0.768 | 70 | 54.5 | 58.3 | 66.7 |
| IR0 Validation | Surveillance Bx | 10 | 3 | 7 | 0.714 | 33.3 | 100 | 100 | 77.8 |
| IR1&IRx Training | Surveillance Bx | 21 | 7 | 14 | 0.791 | 85.7 | 64.3 | 54.5 | 90 |
| IR1&IRx Validation | Surveillance Bx | 12 | 5 | 7 | 0.971 | 100 | 85.7 | 83.3 | 100 |