A modeling approach for mean fluorescence intensity value harmonization and cutoff prediction for luminex single antigen bead assays of two different vendors

Karahan, G.E.; Haasnoot, G.W.; Voogt-Bakker, K.; Claas, F.H.J.; Roelen, D.; Heidt, S.

Received: 6 February 2023 | Revised: 28 March 2023 | Accepted: 18 April 2023

DOI: 10.1111/tan.15082

ORIGINAL ARTICLE

HLA Immune Response Genetics WILEY

A modeling approach for mean fluorescence intensity value harmonization and cutoff prediction for luminex single antigen bead assays of two different vendors

Gonca E. Karahan | **Geert W. Haasnoot** | **Kim Voogt-Bakker** | **Frans H. J. Claas** | **Dave Roelen** | **Sebastiaan Heidt**

Department of Immunology, Leiden University Medical Center, Leiden, The Netherlands

Correspondence
Gonca E. Karahan, Department of Immunology, Leiden University Medical Center, Leiden, The Netherlands.
Email: g.e.karahan@lumc.nl

Luminex single antigen bead (SAB) kits from One Lambda (OL) and Lifecodes (LC) are widely used for HLA antibody detection but have substantial differences in design and assay protocol resulting in different mean fluorescence intensity (MFI) values. Here, we present a non-linear modeling approach to accurately convert MFI values between two vendors and to establish user-independent MFI cutoffs when analyzing big datasets. HLA antibody data from a total of 47 EDTA-treated sera tested using both OL and LC SAB kits were analyzed. MFI comparisons were made for the common 84 HLA class I and 63 class II beads. In the exploration set (n = 24), a non-linear hyperbola model on raw MFI corrected by locus-specific highest self MFI subtraction yielded the highest correlation (class I r^2: 0.946, class II r^2: 0.898). Performance of the model was verified in an independent validation set (n = 12) (class I r^2: 0.952, class II r^2: 0.911). Furthermore, in an independent cohort of post-transplant serum samples (n = 11) using the vendor-specific MFI cutoffs dictated by the current model, we found 94% accuracy in bead-specific reactivity assignments by the two vendors. We recommend using the non-linear hyperbola modeling approach with self HLA correction and locus-specific analyzes to harmonize MFI values between two vendors in particular research datasets. As there are considerable variations between the two assays, using MFI conversion for individual patient samples is not recommended.

KEYWORDS
HLA antibody, Luminex, MFI cutoff, single antigen bead assay

Abbreviations: ABMR, antibody-mediated rejection; BCM, background corrected mean fluorescence intensity; EDTA, ethylenediaminetetraacetic acid; LC, lifecodes/immucor; MFI, mean fluorescence intensity; NGS, next-generation sequencing; OL, One Lambda; SAB, single antigen bead.

1 | INTRODUCTION

Luminex single antigen bead (SAB) assays serve as excellent tools to define HLA antibody specificities owing to

their high specificity and sensitivity. Currently, SAB assays are commercialized by two vendors, One Lambda (OL) and Lifecodes-Immucor (LC). Both vendors provide mean fluorescence intensity (MFI) values as the readout for antibody binding to the SAB. Despite being a semi-quantitative readout, MFI values are frequently used as a cutoff for clinical decision making.[1] However, differences in manufacturing process (e.g., antigen density and integrity coated on beads) and test protocol of SAB assays between two vendors can lead to variances in MFI values. Therefore, utilization of haphazardly selected MFI values as cutoffs, using a single cutoff for both class I and class II, or using a generic cutoff based on the experience with only one vendor may lead to misinterpretation of SAB results.[2–4]

Thus far, the two vendors' SAB results have been compared in a few studies in which either an MFI range resulting in best agreement between two vendors was explored or a single cutoff for both vendors was used to assign positives.[5–7] While the optimum MFI cutoff for best agreement between two vendors was proposed to be between 1000 and 1500 MFI by Reed et al.[7] Clerkin et al. showed an increase in agreement between two vendors when the MFI cutoff for positive antibody assignment was 4000 instead of 1000.[6] Moreover, authors used linear regression models in these studies to assess the correlation of MFI values between two vendors. Noteworthy, there have been no studies providing a strategy to model the relationship of two vendors' MFI values enabling conversion of MFI from one vendor to another.

Many centers use kits from both vendors simultaneously or switch from one to another, necessitating harmonization of the MFI values from two vendors. In the current study, we aimed to generate a method to apply when harmonizing cutoffs between MFI values of two Luminex SAB vendors in big datasets of HLA antibody results is required. To this aim, we first developed a non-linear hyperbola model to better understand how MFI values from two different vendors behave with respect to each other in a given dataset and then used this model to establish data-driven, comparable, vendor-specific cutoffs completely independent of users' experience with one or the other vendor.

2 | MATERIALS AND METHODS

2.1 | HLA typing and antibody data recruitment

HLA antibody and typing data from HLA-immunized individuals were retrieved from the local HLA laboratory, Leiden, the Netherlands with informed consent for the use of their data. HLA antibody and typing data from a total 47 HLA-sensitized individuals were included in the study. HLA antibody data from serum samples of pregnancy-immunized women were analyzed in exploration (n = 24) and validation sets (n = 12). For further validation, as well as for determining clinical utility, post-transplant HLA antibody data from sensitized kidney transplant patients (n = 11) were analyzed.

HLA antibody data were generated by different technicians using various lots of Luminex SAB kits within a 3-year time period (Table S1). Luminex HLA class I and II SAB kits from both OL (LabScreen, One Lambda Inc., Canoga Park, CA, USA) and LC (LSA, Immucor Transplant Diagnostics, Stamford, CT, USA) were used upon pre-treatment of serum with ethylenediaminetetraacetic acid (EDTA), at a final concentration of 8 mM. For LC SAB test, 7.5 μL serum was mixed with 30 μL class I or II beads and tested by a previously described protocol using 75% of reagents.[8] Data were analyzed using MATCH IT! antibody software version 1.3.1 (Immucor). Results were expressed as raw MFI or background-corrected MFI (BCM). For the OL SAB test, 20 μL serum was mixed with 4 μL class I or II beads and further tested as recommended by the manufacturer. Data were analyzed using HLA FUSION antibody software version 3.4.18 (One Lambda). Results were expressed as raw or baseline (normalized) MFI.

HLA typing data of pregnancy immunized serum donors were generated by complement-dependent cytotoxicity method for class I, and by low resolution polymerase chain reaction using sequence-specific oligonucleotides for class II, as described elsewhere.[9,10] In cases where self HLA typing was missing for a particular locus, typing was completed based on linkage with other loci where possible. HLA typing data of all patients and donors in the transplantation cohort were generated by next–generation sequencing (NGS) for HLA–A, B, C, DRB1, DRB3/4/5, DQB1, DQA1, DPB1, and DPA1 loci on Illumina platform (Illumina, San Diego, CA) using NGSgo kits (GenDx, Utrecht, the Netherlands), as previously described,[11] allowing for accurate HLA mismatch identification between the patients and donors.

2.2 | Model construction and statistical analyzes

For determining the relationship between the MFI values of two vendors, regression analysis was used. Two approaches were investigated: linear and hyperbola regression.

The r-squared (r^2) was used as a measure of goodness of fit of the models. This described how close the data points were to the fitted curve that was subsequently

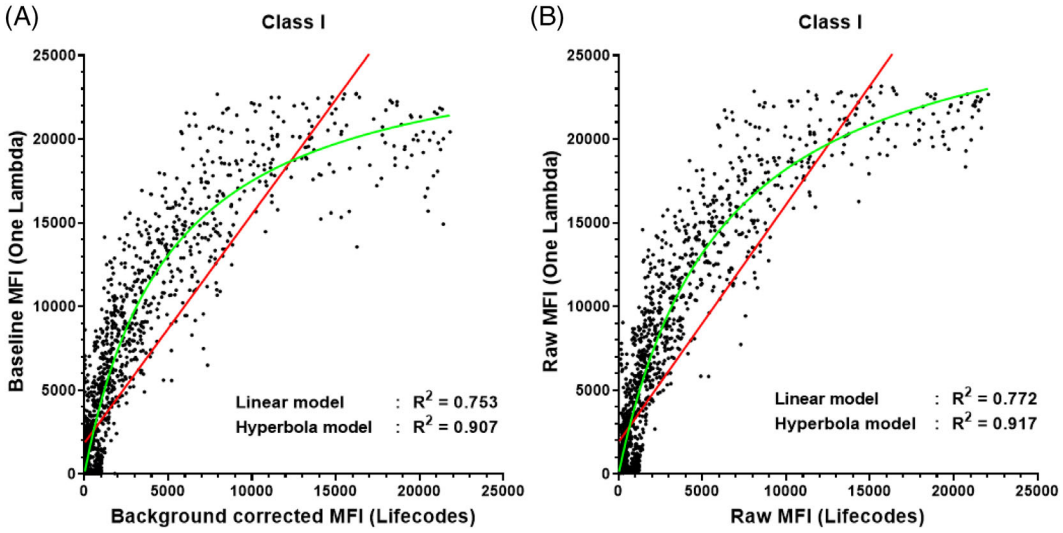

FIGURE 1 Non-linear model optimally fits the data. Linear and non-linear regression models for common HLA class I beads on (A) background corrected/baseline MFI; (B) raw MFI. The linear model is depicted in red and the non-linear hyperbola model is depicted in green.

used for interpolation. Interpolations enabled finding the equivalent MFI in one vendor for a theoretical MFI in the other vendor's test. Beads outside the 95% prediction intervals where defined as outliers.

Several MFI cutoffs for one vendor with the interpolated MFI from the other vendor were used to investigate the agreement in the assignment of positive or negative beads. The level of agreement was analyzed with the use of sensitivity, specificity, accuracy, and the phi coefficient.

Scatter plots were used to visualize the data and the models with linear and log-2 scales. Analyzes were performed with Graphpad Prism (version 9.0.1).

3 | RESULTS

3.1 | Non-linear hyperbola model with correction for self HLA fits the data better than a linear model

To determine the optimal model with the highest goodness of fit, we compared linear and non-linear regression models, applied on background adjusted MFI values of 84 common HLA class I beads in the exploration cohort (n = 24).

Whereas the linear modeling approach on background adjusted MFI values (baseline MFI in OL and BCM in Lifecodes) resulted in an $r^2 = 0.753$, the non-linear hyperbola model led to an $r^2 = 0.907$ (Figure 1A). Each vendor has their own way of background adjustment on raw MFI values, which may introduce bias. Therefore, we applied the models on raw MFI values. Figure 1B shows that application of the non-linear hyperbola model on raw MFI values further increased the goodness of fit to an r^2 value of 0.917. Therefore, all further analyzes were performed on raw MFI values from both vendors.

As we aimed our model to be applicable to all samples regardless of possible background noise, we next applied background correction for self HLA by subtracting the highest locus-specific self HLA bead MFI from raw MFI. In case of missing locus-specific HLA typing, we used the bead with the lowest locus-specific MFI for correction. This background correction not only resolved the background noise for HLA class I (Figure S1A), but also led to an improvement in r^2 values reaching 0.946 for HLA class I (Figure S1B), (Table 1). Interestingly, when we performed a locus-specific analysis using the self HLA-corrected non-linear hyperbola model, we especially found a major improvement for HLA-C. Whereas for HLA-A ($r^2 = 0.907$) and HLA-B ($r^2 = 0.935$) loci, self HLA correction led to a minor improvement (HLA-A $r^2 = 0.933$ and HLA-B $r^2 = 0.953$), r^2 values for HLA-C increased from 0.684 to 0.889 upon self HLA correction (Table 1).

For HLA class II, we analyzed 63 common HLA class II-coated beads between two vendors. Analysis using the non-linear hyperbola model with or without self-HLA correction resulted in $r^2 = 0.898$ in comparison to $r^2 = 0.843$ by the linear model. Locus-specific analyzes using the self HLA-corrected non-linear hyperbola model revealed r^2 value of 0.945 for HLA-DR, 0.894 for HLA-DQ and 0.944 for HLA-DP (Table 1).

To verify these results, we tested both the linear and non-linear modeling approaches on an independent set

TABLE 1 Overview of all models used in the exploration and validation sets including r^2 values and interpolated MFI values.

			Exploration Cohort						Validation Cohort					
					Interpolated MFI OL		Interpolated MFI LC				Interpolated MFI OL		Interpolated MFI LC	
HLA locus	**Correction**	**Model**	**N**	**R^2**	**LC = 1000**	**LC = 3000**	**OL = 3000**	**OL = 8000**	**N**	**R^2**	**LC = 1000**	**LC = 3000**	**OL = 3000**	**OL = 8000**
HLA Class I	Raw MFI	Linear	2016	0.772	3322	6160	773	4297	1008	0.764	3446	6914	743	3626
	Raw MFI	Hyperbola		0.917	4110	9637	699	2300		0.934	5023	11,167	550	1821
	Raw MFI & Self-HLA	Hyperbola		0.946	5044	10,830	540	1846		0.952	6158	12,280	415	1445
HLA-A	Raw MFI	Linear	672	0.737	3444	6280	687	4214	336	0.706	3474	7259	750	3392
	Raw MFI	Hyperbola		0.907	4129	9730	696	2275		0.925	5015	11,429	556	1796
	Raw MFI & Self-HLA	Hyperbola		0.933	4942	10,768	556	1890		0.944	5995	12,485	439	1474
HLA-B	Raw MFI	Linear	1032	0.791	3568	6323	588	4217	516	0.780	3713	7025	570	3589
	Raw MFI	Hyperbola		0.935	4179	9694	685	2271		0.942	5109	11,213	537	1794
	Raw MFI & Self-HLA	Hyperbola		0.953	5010	10,759	545	1874		0.958	6074	12,170	423	1471
HLA-C	Raw MFI	Linear	312	0.652	2697	6544	1158	3757	156	0.825	2659	6272	1189	3956
	Raw MFI	Hyperbola		0.684	3524	9021	839	2578		0.917	4464	10,305	635	2083
	Raw MFI & Self-HLA	Hyperbola		0.889	6647	12,505	366	1317		0.947	7574	12,981	291	1093
HLA Class II	Raw MFI	Linear	1512	0.843	1933	4690	1774	5402	756	0.798	2081	5008	1628	5045
	Raw MFI	Hyperbola		0.898	2656	6855	1143	3663		0.901	3491	8322	842	2835
	Raw MFI & Self-HLA	Hyperbola		0.898	2756	7016	1098	3568		0.911	3943	8892	726	2547
HLA-DR	Raw MFI	Linear	768	0.912	1789	4374	1937	5805	384	0.896	1889	4515	1846	5654
	Raw MFI	Hyperbola		0.943	2146	5808	1431	4420		0.963	2757	7022	1098	3563
	Raw MFI & Self-HLA	Hyperbola		0.945	2141	5785	1435	4447		0.968	2895	7224	1041	3447
HLA-DQ	Raw MFI	Linear	408	0.740	2399	5427	1397	4699	204	0.635	2262	5183	1505	4929
	Raw MFI	Hyperbola		0.876	4105	9461	697	2344		0.802	4758	10,044	574	2055
	Raw MFI & Self-HLA	Hyperbola		0.894	5079	10,664	531	1869		0.846	6392	11,530	372	1430

TABLE 1 (Continued)

HLA locus	Correction	Model	Exploration Cohort				Validation Cohort			
			N	R^2	Interpolated MFI OL		Interpolated MFI LC		N	R^2	Interpolated MFI OL		Interpolated MFI LC	
					LC = 1000	LC = 3000	OL = 3000	OL = 8000			LC = 1000	LC = 3000	OL = 3000	OL = 8000
HLA-DP	Raw MFI	Linear	336	0.937	1816	5057	1731	4817	168	0.954	2339	6747	1300	3568
	Raw MFI	Hyperbola		0.943	1996	5697	1523	4351		0.975	3228	8417	924	2817
	Raw MFI & Self-HLA	Hyperbola		0.944	2005	5710	1517	4346		0.979	3418	8660	867	2709

Abbreviations: LC, Lifecodes; MFI, Mean fluorescence intensity; OL, One Lambda.

of samples from pregnancy-immunized women (validation cohort, n = 12). As shown in Table 1, the non-linear hyperbola model again yielded higher r^2 values for both HLA class I (0.934) and HLA class II (0.901) when compared to the linear model (HLA class I $r^2 = 0.764$, HLA class II $r^2 = 0.798$). Self HLA correction in combination with the hyperbola model resulted in even higher r^2 values (HLA class I: 0.952 and HLA class II: 0.911).

3.2 | Interpolations: From One Lambda to Lifecodes and vice versa

Having achieved highest r^2 values for both HLA class I and class II using the hyperbola model with self-HLA correction on raw MFI, we then performed interpolations from 1000 and 3000 MFI LC to OL as well as from 3000 and 8000 OL MFI to LC (Table 1).

Interpolations from self HLA-corrected raw MFI of 1000 in LC corresponded to higher OL MFI values for HLA class I (5044) and HLA class II (2756) in the exploration as well as validation sets (HLA class I: 6158 and HLA class II: 3943).

Locus-specific interpolations from 1000 LC MFI in the exploration set revealed that OL MFI values for HLA-A (4942) and HLA-B (5010) were relatively close to each other while HLA-C was remarkably different (6647). Within HLA class II, 1000 MFI in LC corresponded to closer values in OL for HLA-DR (2141) and HLA-DP (2005) whereas interpolated MFI for HLA-DQ (5079) was rather discrepant. Similarly, interpolations from 1000 LC MFI to OL in the validation cohort resulted in HLA-C locus corresponding to the highest MFI (7574) whereas MFI values for HLA-A (5995) and HLA-B (6074) loci were relatively close to each other. Interpolations from 3000 LC MFI resulted in 10,830 and 12,280 OL MFI for HLA class I in exploration and validation sets, respectively, while corresponding to 7016 and 8892 for HLA class II. Locus-specific interpolations from 3000 MFI in LC revealed similar trends to interpolations from 1000 LC MFI in both exploration and validation sets.

When interpolations from 3000 OL MFI were performed, the corresponding MFI value in LC was lower for HLA class I (540) than HLA class II (1098) in the exploration as well as validation set (HLA class I MFI: 415, HLA class II MFI: 726). Likewise, interpolations from 8000 OL MFI resulted in 1846 and 1445 LC MFI for HLA class I in exploration and validation set, respectively, while corresponding to 3568 and 2547 for HLA class II (Table 1). Further locus-specific interpolations from OL to LC in both exploration and the validation sets revealed a similar picture to interpolations from LC to OL regarding HLA-C within class I and HLA-DQ within HLA class II being the most discrepant loci.

3.3 | Divergent beads: Are there any?

Considering the significant differences in interpolated MFI within different loci, we applied locus-specific cutoffs based on interpolations and assessed the sensitivity, specificity and accuracy of the assignments between two vendors. Overall, when LC (1000 and 3000 MFI) and interpolated OL MFI values were used as cutoffs, median specificity, sensitivity and accuracy were 98.0% (range:95.5%–99.6%), 92.2% (range: 68.6%–94.9%) and 96.5% (range: 93.5%–98.5%) in exploration and 98.7% (range: 96.4%–100%), 92.3% (66.7%–100%) and 98.3% (range: 94.0%–99.4%) in validation sets, respectively. Likewise when OL (3000 and 8000 MFI) and interpolated LC MFI were established as cutoffs in exploration set, a median specificity of 97.8% (range: 94.7–99.7%), sensitivity of 93.4% (range: 72.5%–100%) and accuracy of 96.6% (92.3%–99.7%) were found. Similarly in validation set, a median specificity of 99.2% (range: 95.9%–100%), sensitivity of 91.9% (range: 86.1%–100%) and accuracy of 97.2% (94.8%–99.4%) were found. Moreover, concordance between the two vendors was also confirmed by the median phi coefficient (rφ) of 0.906 (range: 0.681–0.988) in all groups in both directions (Table 2A and 2B). In addition, despite resulting in very low (<500) interpolated LC MFI values, we also included 1000 OL MFI (Table 2B) to this analysis as it is a commonly used cutoff in many HLA laboratories. While median sensitivity, specificity and accuracy were comparable between exploration (98.0%, 92.0% and 94.6%) and validation cohorts (98.4%, 89.6% and 94.5%) at this cutoff, sensitivity tended to be the lowest at 1000 OL MFI and interpolated MFI cutoff in comparison to higher cutoff values for HLA-C and HLA-DQ both in exploration (67.1% and 89.6%) and validation (79.5% and 83.7%) sets.

Lastly, we investigated whether there were outlier beads that were outside the 95% prediction intervals. To do this, for a given MFI value in one vendor (LC: 1000 or OL: 3000), we compared the observed MFI with the expected (predicted) MFI in the other vendor. As a result, we found several beads multiple times outside the prediction intervals. However, only a few were exclusively above or below the expected values occurring ≥2 times in both cohorts, as shown in Table S2A. These were HLA-class I beads coated with *B*42:01*, *B*46:01*, *B*47:01*, *B*48:01*, and *C*05:01* and HLA-class II beads coated with *DRB1*01:01*, *DQB1*05:01-DQA1*01:01* and *DQB1*06:01-DQA1*01:03*. Remarkably, not all discrepancies in MFIs of above-mentioned beads resulted in disagreement of the assignments between the two vendors (Tables S2B and S2C), suggesting adequate overlap between vendors. There were no beads negative in both exploration and validation cohorts, indicating the complete representation of all specificities among common beads (data not shown).

3.4 | Agreement in bead-specific reactivity assignment using a transplantation cohort

As a second validation step, we applied the non-linear hyperbola model with self HLA correction to the data of a cohort of 11 post-transplant serum samples and again found higher r^2 values for both HLA class I (0.939) and class II (0.925) in comparison to those obtained when linear models (HLA class I $r^2 = 0.817$; HLA class II $r^2 = 0.891$) were used.

Next, we defined the HLA mismatches between patients and donors (Table S3) and compared the bead reactivity specific for each mismatched HLA between SAB kits according to cutoffs based on 1000 LC MFI and locus-specific interpolations to OL, as dictated by the non-linear hyperbola model of the transplantation cohort. After excluding donor antigens not included in SAB panels as well as the ones corresponding to beads that were not common between the vendors, bead-specific reactivity for a total of 87 HLA mismatches was assessed. Based on the 1000 MFI LC and the interpolated OL cutoffs, a total of 20 bead-specific reactivity were assigned positive and 62 were assigned negative by both vendors reaching up to an overall agreement of 94%. Disagreement between two vendors were found for 5 bead-specific reactivity (6%), 4 of which were only positive by LC and 1 were only positive by OL. While 2 out of 5 beads in disagreement had a clear MFI discrepancy between the two vendors (*DRB3*01:01* and *DQB1*02:01-DQA1*05:01*), the remaining three beads had borderline positivity either on one vendor's side (*A*01:01*) or by both vendors (*DRB1*07:01* and *C*01:02*) (Table 3, upper panel).

When we used 3000 MFI OL and the interpolated LC cutoffs, an agreement of 93% for a total of 81 beads was found (positive = 21, negative = 60) between two vendors. Disagreement was found for a total of 6 beads (7%), 5 of which were only positive by LC and 1 was only positive by OL. Four out of 6 beads in disagreement at this cutoff (*A*01:01*, *DRB1*07:01*, *DRB3*01:01* and *DQB1*02:01-DQA1*05:01*) were also found to be in disagreement when the above-mentioned 1000 MFI LC and interpolated OL MFI cutoffs were used. Remaining two beads had borderline positivity on one vendor's side (*C*03:04* and *DQB1*05:02-DQA1*01:02*) (Table 3, lower panel).

3.5 | Performance of the original hyperbola model when employed on different cohorts

We next investigated whether the model developed using the dataset in the exploration cohort could be used on another dataset generated using the same Luminex SAB

TABLE 2A Quality of bead-specific reactivity assignments when interpolated MFI values are used as locus-specific cutoffs (from Lifecodes to One Lambda).

MFI LC	MFI OL	Locus	LC/OL neg/neg	LC/OL pos/neg	LC/OL neg/pos	LC/OL pos/pos	Specificity (%)	Sensitivity (%)	Accuracy (%)	r_φ	N
Exploration Cohort											
1000	4942	HLA-A	408	16	13	235	96.2	94.8	95.7	0.908	672
3000	10,768	HLA-A	515	24	20	113	95.5	85.0	93.5	0.796	672
1000	5010	HLA-B	588	16	22	406	97.4	94.9	96.3	0.924	1032
3000	10,759	HLA-B	741	22	14	255	97.1	94.8	96.5	0.911	1032
1000	6647	HLA-C	289	3	3	17	99.0	85.0	98.1	0.840	312
3000	12,505	HLA-C	301	3	2	6	99.0	75.0	98.4	0.699	312
1000	2141	HLA-DR	539	6	17	206	98.9	92.4	97.0	0.927	768
3000	5785	HLA-DR	606	11	12	139	98.2	92.1	97.0	0.905	768
1000	5079	HLA-DQ	317	9	6	76	97.2	92.7	96.3	0.887	408
3000	10,664	HLA-DQ	340	8	6	54	97.7	90.0	96.6	0.865	408
1000	2005	HLA-DP	280	1	4	51	99.6	92.7	98.5	0.945	336
3000	5710	HLA-DP	296	5	11	24	98.3	68.6	95.2	0.728	336
Validation Cohort											
1000	5995	HLA-A	243	9	11	73	96.4	86.9	94.0	0.840	336
3000	12,485	HLA-A	298	4	1	33	98.7	97.1	98.5	0.922	336
1000	6074	HLA-B	327	9	10	170	97.3	94.4	96.3	0.919	516
3000	12,170	HLA-B	413	6	10	87	98.6	89.7	96.9	0.897	516
1000	7574	HLA-C	131	0	2	23	100.0	92.0	98.7	0.952	156
3000	12,981	HLA-C	137	3	0	16	97.9	100.0	98.1	0.908	156
1000	2895	HLA-DR	317	4	9	54	98.8	85.7	96.6	0.874	384
3000	7224	HLA-DR	346	2	4	32	99.4	88.9	98.4	0.906	384
1000	6392	HLA-DQ	177	0	2	25	100.0	92.6	99.0	0.957	204
3000	11,530	HLA-DQ	182	4	6	12	97.8	66.7	95.1	0.681	204
1000	3418	HLA-DP	134	1	0	33	99.3	100.0	99.4	0.982	168
3000	8660	HLA-DP	147	1	1	19	99.3	95.0	98.8	0.943	168

TABLE 2B Quality of bead-specific reactivity assignments when interpolated MFI values are used as locus-specific cutoffs (from One Lambda to Lifecodes).

MFI OL	MFI LC	Locus	LC/OL neg/neg	LC/OL pos/neg	LC/OL neg/pos	LC/OL pos/pos	Specificity (%)	Sensitivity (%)	Accuracy (%)	r_φ	N
Exploration Cohort											
1000	171	HLA-A	318	6	36	312	98.1	89.7	93.8	0.879	672
3000	556	HLA-A	384	6	8	274	98.5	97.2	97.9	0.957	672
8000	1890	HLA-A	465	26	26	155	94.7	85.6	92.3	0.803	672
1000	167	HLA-B	408	13	35	576	96.9	94.3	95.3	0.905	1032
3000	545	HLA-B	539	12	20	461	97.8	95.8	96.9	0.938	1032
8000	1874	HLA-B	680	19	19	314	97.3	94.3	96.3	0.916	1032
1000	111	HLA-C	230	6	25	51	97.5	67.1	90.1	0.717	312
3000	336	HLA-C	269	3	11	29	98.9	72.5	95.5	0.787	312
8000	1317	HLA-C	295	3	2	12	99.0	85.7	98.4	0.820	312
1000	453	HLA-DR	498	11	15	244	97.8	94.2	96.6	0.924	768
3000	1435	HLA-DR	560	13	19	176	97.7	90.3	95.8	0.889	768
8000	4447	HLA-DR	637	14	5	112	97.8	95.7	97.5	0.908	768
1000	162	HLA-DQ	251	32	13	112	88.7	89.6	89.0	0.755	408
3000	531	HLA-DQ	301	11	8	88	96.5	91.7	95.3	0.872	408
8000	1869	HLA-DQ	331	10	5	62	97.1	92.5	96.3	0.871	408
1000	492	HLA-DP	253	4	19	60	98.4	75.9	93.2	0.803	336
3000	1517	HLA-DP	286	1	0	49	99.7	100.0	99.7	0.988	336
8000	4346	HLA-DP	317	3	0	16	99.1	100.0	99.1	0.913	336
Validation Cohort											
1000	135	HLA-A	172	6	22	136	96.6	86.1	91.7	0.836	336
3000	439	HLA-A	219	8	8	101	96.5	92.7	95.2	0.891	336
8000	1474	HLA-A	262	9	7	58	96.7	89.2	95.2	0.849	336
1000	129	HLA-B	240	1	19	256	99.6	93.1	96.1	0.925	516
3000	423	HLA-B	280	5	17	214	98.2	92.6	95.7	0.914	516
8000	1471	HLA-B	349	15	12	140	95.9	92.1	94.8	0.875	516
1000	87	HLA-C	110	2	9	35	98.2	79.5	92.9	0.823	156
3000	291	HLA-C	120	0	5	31	100.0	86.1	96.8	0.909	156
8000	1093	HLA-C	131	1	2	22	99.2	91.7	98.1	0.925	156
1000	322	HLA-DR	272	2	7	103	99.3	93.6	97.7	0.942	384

test protocol. To explore the performance of the model on independent datasets, we applied the non-linear hyperbola model developed in the exploration set onto the validation and transplantation sets. As shown in Figure 2A,B, applying the hyperbola model of the exploration set on validation set yielded r^2 values of 0.940 and 0.892, for HLA class I and class II, respectively. These values were similar to the original validation set model applied on the validation set itself (class I $r^2 = 0.952$ and class II $r^2 = 0.911$), indicating consistency of the model. Moreover, by applying the original model developed using the exploration set to transplantation cohort, we were able to confirm the consistency of the model, as depicted by r^2 values of 0.932 and 0.900 for HLA class I and class II, respectively, which were almost identical to r^2 values the original model of transplantation dataset (class I $r^2 = 0.939$ and class II $r^2 = 0.925$) (Figure 2C,D). This finding was further explored by using cutoff values generated in exploration and validation cohorts to asses bead-specific reactivity assignments in transplantation cohort. As shown in Table 3, assignments overlapped for all bead-specific reactivities except for *DRB1*07:01* bead which had borderline positivity.

4 | DISCUSSION

Interpretation of luminex SAB test results partially relies on assessment of MFI values in combination with patient-donor specific information.[1,3,12] Two vendors provide several outputs in their analysis software deriving from employment of different means of corrections on raw MFI values.[13] These outputs include baseline MFI and normalized background (NBG) ratio in HLA FUSION (OL) and BCM, antigen density-corrected BCM (AD-BCM) as well as ratios of raw MFI or BCM to the lowest ranked antigen values of the SAB panel in MATCH IT! (LC). These parameters can be used on their own or in combination to assign an antibody positive or negative.[14]

So far, few studies compared the performance of SAB kits from the two vendors.[5,6] In these studies, linear regression was used for modeling the relationship between baseline MFI in OL and BCM in LC with correlations reaching up to a maximum of $r^2 = 0.693$. Here, using a non-linear hyperbola model on baseline/background-corrected as well as raw MFI for HLA class I, we show a significantly higher correlation with r^2 values reaching up to = 0.907 and 0.917, respectively, clearly revealing the non-linear relationship of MFI values between the two vendors' output.

Although vendors have greatly improved their SAB kits over the years, interpretation of SAB assay results

TABLE 2B (Continued)

MFI OL	MFI LC	Locus	LC/OL neg/neg	LC/OL pos/neg	LC/OL neg/pos	LC/OL pos/pos	Specificity (%)	Sensitivity (%)	Accuracy (%)	r_φ	N
3000	1041	HLA-DR	320	3	7	54	99.1	88.5	97.4	0.901	384
8000	3447	HLA-DR	351	2	1	30	99.4	96.8	99.2	0.948	384
1000	110	HLA-DQ	146	15	7	36	90.7	83.7	89.2	0.701	204
3000	372	HLA-DQ	167	5	1	31	97.1	96.9	97.1	0.896	204
8000	1430	HLA-DQ	178	1	3	22	99.4	88.0	98.0	0.907	204
1000	273	HLA-DP	129	2	1	36	98.5	97.3	98.2	0.949	168
3000	867	HLA-DP	133	1	0	34	99.3	100.0	99.4	0.982	168
8000	2709	HLA-DP	144	1	3	20	99.3	87.0	97.6	0.897	168

Note: Accuracy depicts the number of true results with the order of LC/OL (pos/pos + neg/neg) divided by the total number of results (pos/pos + pos/neg + neg/pos + neg/neg).
Abbreviations: LC, Lifecodes; neg, negative; OL, One Lambda; pos, positive.

TABLE 3 Disagreements in bead-specific reactivity assignments corresponding to mismatched donor HLA in transplantation cohort.

		MFI observed		Cut-off				Bead-specific reactivity assignment			
Patient #	Bead	LC	OL	LC	int OL (tx)	int OL (exp)	int OL (val)	LC	OL (tx)	OL (exp)	OL (val)
Cutoff: Lifecodes (1000 MFI) and interpolated OL MFI											
7	*A*01:01*	1121	223	1000	3479	4942	5994	pos	neg	neg	neg
4	*C*01:02*	1260	6122	1000	6403	6646	7574	pos	neg	neg	neg
3	*DRB1*07:01*	1028	2328	1000	3073	2141	2895	pos	neg	pos	neg
9	*DRB3*01:01*	3529	32	1000	3073	2141	2895	pos	neg	neg	neg
7	*DQB1*02:01-DQA1*05:01*	293	13,775	1000	4083	5078	6392	neg	pos	pos	pos
Cutoff: One Lambda (3000 MFI) and interpolated LC MFI											
7	*A*01:01*	1121	223	3000	621	556	439	neg	pos	pos	pos
11	*C*03:04*	892	2952	3000	331	366	291	neg	pos	pos	pos
3	*DRB1*07:01*	1028	2328	3000	738	1435	1041	neg	pos	neg	neg
9	*DRB3*01:01*	3529	32	3000	738	1435	1041	neg	pos	pos	pos
7	*DQB1*02:01-DQA1*05:01*	293	13,775	3000	662	531	372	pos	neg	neg	neg
7	*DQB1*05:02-DQA1*01:02*	807	32	3000	662	531	372	neg	pos	pos	pos

Note: Common beads of disagreement at different cutoffs are underlined. Assignments in transplantation cohort were done according to the cutoffs dictated by the model generated within the transplantation (tx) dataset itself as well as using cutoffs dictated by the models generated in exploration (exp) and validation (val) datasets.

Abbreviations: int, interpolated; LC, Lifecodes; MFI, Mean fluorescence intensity; neg, negative; OL, One Lambda; pos, positive.

can still be challenging due to technical problems such as complement interference and prozone effect.[15–18] In a recent study comparing luminex SAB kits from the two vendors, 11 out 125 serum samples analyzed were found to be responsible for 80% of the outliers between the two SAB assays.[6] Noteworthy, no correlation ($r^2 = 0.0008$) for class I and a very weak correlation ($r^2 = 0.063$) for class II were found in these samples at initial testing whereas further dilution of the samples in OL kit improved both class I ($r^2 = 0.693$) and class II ($r^2 = 0.383$) correlations, indicating a prozone effect in OL SAB assay.[6]

While such a prozone effect in OL SAB assay could be attributed to the higher serum/bead ratio in the test protocol, issues related to lot-to lot variability of SAB kits, inter-machine variability, non-specific background noise in samples as well as those related to antigen density and integrity on beads can still complicate analyzes in both vendors.[7,14,19–22] In our study, EDTA treated samples were tested using various lots of SAB kits from both vendors over a 3-year time period, on the same Luminex machine, excluding inter-machine variability. For Lifecodes, we used a 25% reduced reagent assay protocol. Reduction of assay reagents upto 50% have been shown to be comparable to 100% of the reagents by Kamburova et al,[8] suggesting that using a reduced reagent test protocol had no influence on the performance of the model. We did not consider lot-to-lot variability in this study however we paid close attention to the background noise observed in a few samples. These samples had relatively higher negative control bead values and reactivity to self HLA-coated beads.[14,21] Rather than excluding these samples from the analyzes, we applied a correction by subtracting raw MFI of the highest locus-specific self-HLA coated bead from the corresponding locus-specific beads. This approach not only resolved the noise in our dataset and enabled us to analyze all our data, but also led to an improved correlation ($r^2 = 0.946$).

Overall, a good correlation and agreement have been found between two vendors' SAB kits in previous studies.[6,7] By setting one of the vendor's SAB kit as the "truth" for HLA antibody positivity, Reed et al found an agreement of 90% for both HLA class I and class II using receiver operating characteristic (ROC) analysis.[7] Importantly, further locus-specific analyzes revealed the poorest area under the curve for HLA-C and HLA-DQ. In the current study, we found lower r^2 values for HLA class II in comparison to class I in all datasets analyzed. In addition, interpolated MFI values for HLA class II were found to be closer to each other between OL and LC than those of class I. Within HLA class II, HLA-DQ was the most discrepant locus with the lowest correlation coefficients in both exploration ($r^2 = 0.894$) and validation ($r^2 = 0.846$) cohorts. These results led us to utilize locus-specific cutoffs for assessment of antibody assignments between the vendors.

MFI values lower than 3000–4000 are commonly reported to have less agreement in antibody

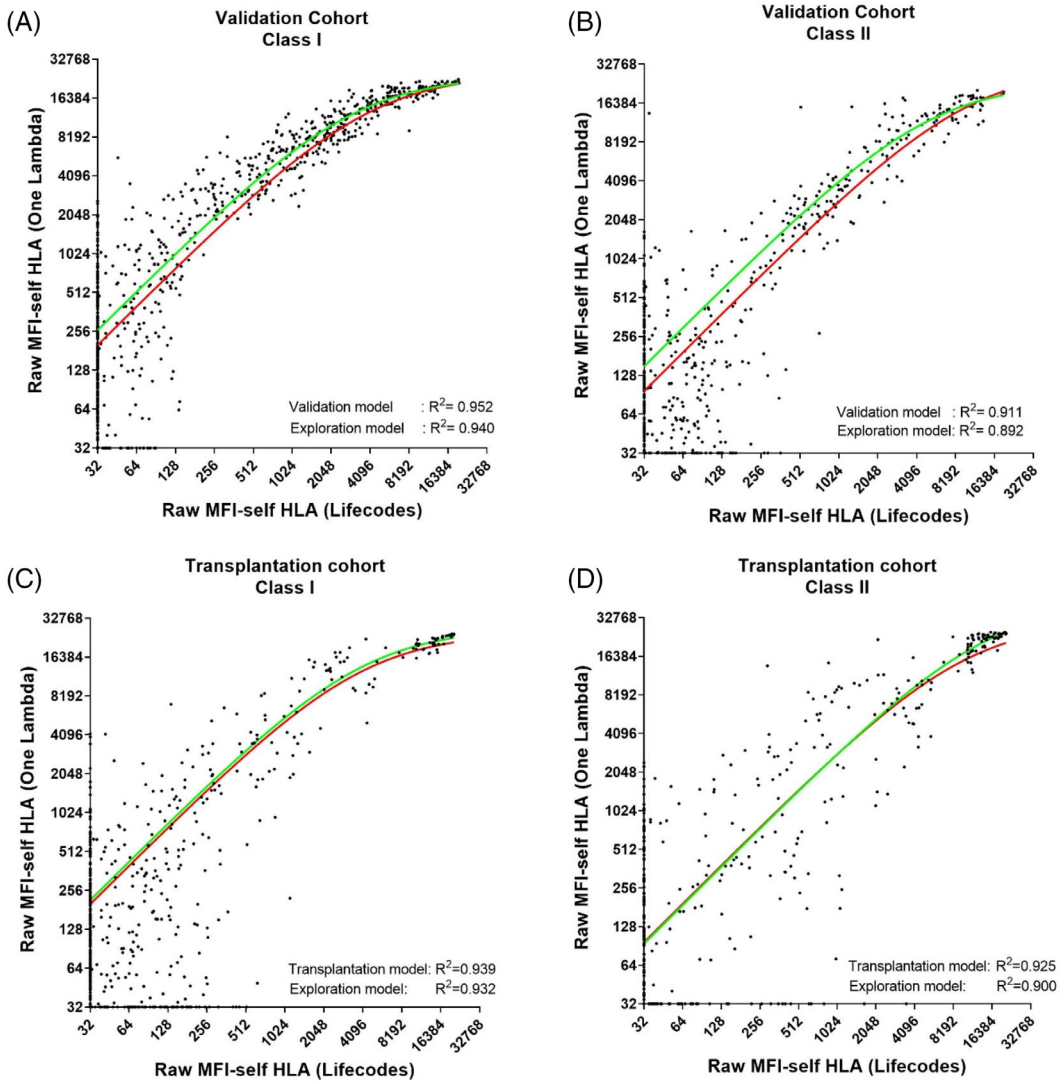

FIGURE 2 Performance of the non-linear hyperbola model on raw MFI values generated in the exploration cohort on independent cohorts. The hyperbola model of the exploration cohort is depicted in red and the hyperbola models of validation and transplantation cohorts are depicted in green. When HLA class I and class II-specific models of exploration cohort (red line) are employed onto independent cohorts such as validation (A: class I and B: class II) or transplantation (C: class I and D: class II) as depicted by the green lines, the model of the exploration set (red line) performs almost as good as the original models of validation and transplantation datasets.

assignments.[4,6,7] In a previous study, an MFI cutoff ranging between 1000 and 1500 has been shown to result in an accuracy varying from 86% to 93% for both class I and class II, suggesting an excellent agreement between the two kits.[7] However, further stratification up to 3000 MFI revealed larger variability for MFI distribution between the two vendors. In the current study, we established cutoffs using interpolated MFI values from LC (1000 and 3000) and OL (3000 and 8000). At all cutoff values applied in both directions for both HLA class I and class II, we found an excellent median specificity (99%), sensitivity (93%) and accuracy (98%), suggesting that our current model for establishing cutoffs was effective not only at higher MFI values but also at intermediate and low MFI where decision making is complicated.

In addition, we also determined outliers consistently present in both exploration and validation sets and found only few beads with MFI values outside the 95% prediction interval.

In the current study, we chose to present interpolations from 1000 and 3000 MFI in Lifecodes as well as 1000, 3000 and 8000 MFI in One Lambda in both exploration and validation cohorts. Mathematical formulas used in our non-linear hyperbola modeling approach as well as an "interpolation calculator" allowing for conversion of different MFI values than the ones presented here between two vendors are provided as a separate Supplementary file (Data S2).

Determination of clinically relevant DSAs is the ultimate goal in HLA antibody detection.[23] In a cohort of 100 post-transplant sera from kidney transplant samples, Bertrand et al compared the performance of OL and LC SAB kits for predicting ABMR.[5] Using their usual cutoff of 500 MFI in OL also for LC, authors found a better correlation for class II ($r = 0.80$) in comparison to class I ($r = 0.63$) for DSA assignment. This contrasts with our results showing a higher overall correlation for class I in comparison class II. However, one should keep in mind that broadness of sensitization in serum samples, differences in test protocols in addition to the statistical model used for analyzes can have affect these results. In their further analyzes modifying their cutoff, Bertrand and colleagues suggested 2705 baseline MFI in OL and 473 BCM in LC to be the most comparable value for predicting ABMR between two vendors.

We studied HLA antibody data of a cohort of 11 post-transplant serum samples to assess the performance of our model for assignments of bead reactivities specific for each mismatched HLA using the locus-specific 1000 LC and 3000 OL MFI and corresponding interpolated MFI values as cutoffs. The model at each cutoff yielded overall an excellent agreement with 94% and 93% accuracy in DSA assignments between the two vendors, respectively. Among the 5 discrepant beads at 1000 LC MFI and interpolated OL cutoff, 3 beads were at borderline positivity in one or both SAB assays. Significantly high MFI values in one vendor while being undoubtedly negative for the other vendor in the remaining 2 beads (*DRB3*01:01* and *DQB1*02:01-DQA1*01:01*) can result from the differences in the amount of antigen coated on beads, as well as reactivity to cryptic epitopes leading to false positives.[3,24–27] Noteworthy, at 3000 OL MFI and interpolated LC cutoff, 4 of the 6 discrepant beads were the same beads found to be discrepant at 1000 LC MFI cutoff in addition to 2 beads at borderline positivity. Furthermore, application of exploration and validation set-specific cutoffs to the transplantation dataset resulted in exactly the same assignments in all beads except one (*DRB1*07:01*) which was at borderline positivity according to transplant dataset specific cutoffs. These result once again confirm the performance of the model.

To our knowledge, this is the first study to provide a non-linear modeling approach enabling conversion of MFI values between two vendors and establishing user-independent, dataset-specific MFI cutoffs. The current model will help to establish comparable MFI cutoffs for the two different kits, at least on the population level for studying large datasets. While, we strongly suggest antibody pattern analyzes for accurate HLA antibody assignments in individual patient samples, one should bear in mind that variation in the two assays per specificity precludes MFI conversion from one vendor to the other for individual patients.

AUTHOR CONTRIBUTIONS

Gonca E. Karahan designed the study, analyzed the data, wrote the manuscript; Geert W. Haasnoot participated in the research design and analyzed the data; Kim Voogt-Bakker participated in recruitment and analyzes of the data; Frans H. J. Claas and Dave Roelen participated in the research design and analyzed the data; Sebastiaan Heidt participated in the research design, analyzed the data and revised the manuscript.

ACKNOWLEDGMENTS

The authors thank the HLA laboratory, Leiden, the Netherlands for providing the data. This study was supported by National Reference Center for Histocompatibility Testing, Leiden, the Netherlands.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are availableupon reasonable request.

ETHICS STATEMENT

Not applicable.

ORCID

Gonca E. Karahan https://orcid.org/0000-0001-5339-7126
Sebastiaan Heidt https://orcid.org/0000-0002-6700-188X

REFERENCES

1. Sullivan HC, Liwski RS, Bray RA, Gebel HM. The road to HLA antibody evaluation: do not rely on MFI. *Am J Transplant*. 2017;17(6):1455-1461.
2. Battle RK, Henderson L, Phelan PJ, Latham K, Turner DM. A case report-two manufacturers SAB testing kits can reveal different HLA antibody profiles-identifying prozone and denatured antigen. *HLA*. 2020;96(1):76-82.
3. Sullivan HC, Krummey SM, Gebel HM, Bray RA. The utility of second single antigen bead assay: clearing the water or stirring up mud? *Hum Immunol*. 2020;81(12):663-670.
4. Chen M, Zoet Y, Roelen D, et al. Towards uniformity in the definition of acceptable mismatches for highly sensitized patients. *HLA*. 2019;94(2):147-153.
5. Bertrand D, Farce F, Laurent C, et al. Comparison of two Luminex single-antigen bead flow cytometry assays for detection of donor-specific antibodies after renal transplantation. *Transplantation*. 2019;103(3):597-603.
6. Clerkin KJ, See SB, Farr MA, et al. Comparative assessment of anti-HLA antibodies using two commercially available Luminex-based assays. *Transplant Direct*. 2017;3(11):e218.

7. Reed EF, Rao P, Zhang Z, et al. Comprehensive assessment and standardization of solid phase multiplex-bead arrays for the detection of antibodies to HLA. *Am J Transplant*. 2013; 13(7):1859-1870.
8. Kamburova EG, Wisse BW, Joosten I, et al. How can we reduce costs of solid-phase multiplex-bead assays used to determine anti-HLA antibodies? *HLA*. 2016;88(3):110-119.
9. Terasaki PI, McClelland JD. Microdroplet assay of human serum cytotoxins. *Nature*. 1964;204:998-1000.
10. Giphart MJ, Verduijn W. The Eurotransplant HLA-DRB oligonucleotide typing set. *Eur J Immunogenet*. 1991;18(1–2): 57-68.
11. Kramer CSM, Franke-van Dijk MEI, Bakker KH, et al. Generation and reactivity analysis of human recombinant monoclonal antibodies directed against epitopes on HLA-DR. *Am J Transplant*. 2020;20:3341-3353.
12. Sullivan HC, Gebel HM, Bray RA. Understanding solid-phase HLA antibody assays and the value of MFI. *Hum Immunol*. 2017;78(7–8):471-480.
13. Ravindranath MH, Filippone EJ, Mahowald G, et al. Significance of the intraindividual variability of HLA IgG antibodies in renal disease patients observed with different beadsets monitored with two different secondary antibodies on a Luminex platform. *Immunol Res*. 2018;66(5):584-604.
14. Karahan GE, de Vaal Y, Bakker K, Roelen D, Claas FHJ, Heidt S. Comparison of different luminex single antigen bead kits for memory B cell-derived HLA antibody detection. *HLA*. 2021;98(3):200-206.
15. Tait BD, Susal C, Gebel HM, et al. Consensus guidelines on the testing and clinical management issues associated with HLA and non-HLA antibodies in transplantation. *Transplantation*. 2013;95(1):19-47.
16. Guidicelli G, Visentin J, Franchini N, et al. Prevalence, distribution and amplitude of the complement interference phenomenon in single antigen flow beads assays. *HLA*. 2018;91(6):507-513.
17. Tambur AR, Campbell P, Claas FH, et al. Sensitization in transplantation: assessment of risk (STAR) 2017 working group meeting report. *Am J Transplant*. 2018;18(7):1604-1614.
18. Karahan GE, Claas FHJ, Heidt S. Technical challenges and clinical relevance of single antigen bead C1q/C3d testing and IgG subclass analysis of human leukocyte antigen antibodies. *Transpl Int*. 2018;31:1189-1197.
19. Konvalinka A, Tinckam K. Utility of HLA antibody testing in kidney transplantation. *J Am Soc Nephrol*. 2015;26(7):1489-1502.
20. Gandhi MJ, Degoey S, Falbo D, et al. Inter and intra laboratory concordance of HLA antibody results obtained by single antigen bead based assay. *Hum Immunol*. 2013;74(3):310-317.
21. Zerrouki A, Ouadghiri S, Benseffaj N, Razine R, Essakalli M. Reason and resolution of high negative control beads in solid-phase immunoassay. *Exp Clin Transplant*. 2018;16(1):38-43.
22. Battle RK, Rennie TJW, Phelan PJ, Abel AA, McConnell S, Turner DM. Highly sensitised patients awaiting deceased donor renal transplants are disadvantaged by the presence of denatured HLA antibody detected in routine HLA antibody testing. *HLA*. 2022;100(1):24-36.
23. Wisse BW, Kamburova EG, Joosten I, et al. Toward a sensible single-antigen bead cutoff based on kidney graft survival. *Transplantation*. 2019;103(4):789-797.
24. Habets T, Hepkema BG, Kouprie N, et al. The prevalence of antibodies against the HLA-DRB3 protein in kidney transplantation and the correlation with HLA expression. *PLoS One*. 2018;13(9):e0203381.
25. Wehmeier C, Honger G, Schaub S. Caveats of HLA antibody detection by solid-phase assays. *Transpl Int*. 2020;33(1):18-29.
26. Morales-Buenrostro LE, Terasaki PI, Marino-Vazquez LA, Lee JH, El-Awar N, Alberu J. "Natural" human leukocyte antigen antibodies found in nonalloimmunized healthy males. *Transplantation*. 2008;86(8):1111-1115.
27. Park BG, Park Y, Kim BS, Kim YS, Kim HS. False positive class II HLA antibody reaction due to antibodies against denatured HLA might differ between assays: One lambda vs. Immucor. *Annal Lab Med*. 2020;40(5):424-427.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Karahan GE, Haasnoot GW, Voogt-Bakker K, Claas FHJ, Roelen D, Heidt S. A modeling approach for mean fluorescence intensity value harmonization and cutoff prediction for luminex single antigen bead assays of two different vendors. *HLA*. 2023;102(5): 557-569. doi:10.1111/tan.15082