Validity and Reproducibility of Measurement of Islet Autoreactivity by T-Cell Assays in Subjects With Early Type 1 Diabetes

Kevan C. Herold,[1] Barbara Brooks-Worrell,[2] Jerry Palmer,[3] H. Michael Dosch,[4] Mark Peakman,[5] Peter Gottlieb,[6] Helena Reijonen,[7] Sefina Arif,[5] Lisa M. Spain,[8] Clinton Thompson,[1] John M. Lachin,[1] and the Type 1 Diabetes TrialNet Research Group*

OBJECTIVE—Type 1 diabetes results from an immune-mediated destruction of β-cells, likely to be mediated by T lymphocytes, but the sensitivity, specificity, and other measures of validity of existing assays for islet autoreactive T-cells are not well established. Such assays are vital for monitoring responses to interventions that may modulate disease progression.

RESEARCH DESIGN AND METHODS—We studied the ability of cellular assays to discriminate responses in patients with type 1 diabetes and normal control subjects in a randomized blinded study in the U.S. and U.K. We evaluated the reproducibility of these measurements overall and to individual analytes from repeat collections.

RESULTS—Responses in the cellular immunoblot, U.K.-ELISPOT, and T-cell proliferation assays could differentiate patients from control subjects with odds ratios of 21.7, 3.44, and 3.36, respectively, with sensitivity and specificity as high as 74 and 88%. The class II tetramer and U.S. ELISPOT assays performed less well. Despite the significant association of the responses with type 1 diabetes, the reproducibility of the measured responses, both overall and individual analytes, was relatively low. Positive samples from normal control subjects (i.e., false positives) were generally isolated to single assays.

CONCLUSIONS—The cellular immunoblot, U.K.-ELISPOT, and T-cell proliferation assays can distinguish responses from patients with type 1 diabetes and healthy control subjects. The limited reproducibility of the measurements overall and of responses to individual analytes may reflect the difficulty in detection of low frequency of antigen-specific T-cells or variability in their appearance in peripheral blood. *Diabetes* **58:2588–2595, 2009**

Type 1 diabetes is caused by T-cell–mediated destruction of β-cells (1). Despite this understanding, there are few tools to identify and track cells that mediate the disease in humans. Several assays that can distinguish antigen-specific responses in patients from normal control subjects have been reported (2–7); however, some of these have not performed well in larger blinded studies (8–10), and their reproducibility has not been systematically studied in a masked workshop.

A means of monitoring cellular responses involved in type 1 diabetes is needed to understand the action of immune therapies and to effectively apply therapies in the clinical setting (11). A biomarker that responds to an effective therapy could be used to rapidly screen candidate therapies for use in a further study to examine effects on longer-term clinical outcomes, such as preservation of β-cell function. A highly sensitive and specific biomarker might also potentially function as a surrogate for clinical outcomes that take a longer time to observe and require a larger number of subjects to study.

In a previous blinded study, Seyfert-Margolis et al. (8) reported that two different assays that measured T-cell proliferative responses to antigens were able to distinguish responses in subjects with type 1 diabetes from those in healthy normal control subjects. However, that study used a single collection from each subject and could not evaluate the reproducibility of the measurements that is needed to assess their utility in clinical trials. Moreover, only the two assays that used fresh cells showed significant discriminant validity (i.e., ability to distinguish participants with type 1 diabetes from normal control subjects). Therefore, the present study was conducted to assess the discriminant ability of five T-cell assays with fresh blood samples. We also assessed the reproducibility of the measurements from repeat collections in individual subjects, both qualitatively with respect to the classification of each subject (positive vs. negative) and quantitatively for the different analytes used in each assay.

RESEARCH DESIGN AND METHODS

Subjects. Sixty-eight control subjects with type 1 diabetes were enrolled, 35 in North America and 33 in the U.K., along with 96 control subjects without type 1 diabetes, 63 in North America and 33 in the U.K. (Table 1). Collections from North American sites were split and distributed among the North American laboratories; those from the U.K. were assayed by the U.K. laboratory.

A greater number of control subjects were studied to provide adequate numbers of subjects with HLA-DR3 and/or DR4 genotypes (12). Two collec-

TABLE 1
Characteristics of subjects enrolled in North America and the U.K. who contributed an evaluable specimen

	North America		U.K.	
	Diabetes	Normal	Diabetes	Normal
n	35	59	29	31
Age (years)	15.0 ± 3.98	26.5 ± 5.65	24.2 ± 6.6	24.7 ± 5.5
Sex (male)	22 (62.9)	26 (44.1)	19 (65.5)	13 (41.9)
Type 1 diabetes duration (months)	6.4 ± 3.05	—	3.3 ± 3.88	—
Whites (%)	79.4	86.4	96.6	90.3
Collections (%)	66	113	51	58
1 only	35	59	29	31
2 only	31	54	22	27
Days between collections				
2–7	8 (25.8)	18 (33.3)	5 (22.7)	14 (51.9)
8–14	14 (45.2)	15 (27.8)	9 (40.9)	7 (27.9)
15–21	4 (12.9)	9 (16.7)	6 (27.3)	2 (7.4)
21–28	5 (16.1)	12 (22.2)	2 (9.1)	4 (14.8)
Autoantibody positive				
ICA512	23 (65.7)	0 (0)	14 (50)	0 (0)
GAD65	22 (62.9)	0 (0)	20 (71.4)	0 (0)
ICA	25 (71.4)	5 (8.6)	13 (46.4)	0 (0)
One	7 (20)	5 (8.6)	7 (24.1)	0 (0)
Two or more	24 (68.6)	0 (0)	15 (51.7)	0 (0)
HLA type				
DR3	7 (20)	16 (27.1)	4 (13.8)	11 (35.5)
DR4	14 (40)	23 (39.0)	16 (55.2)	17 (54.8)
DR3/4	11 (31.4)	9 (15.3)	9 (31)	3 (9.7)
Neither	3 (8.6)	11 (18.6)	—	—

Data are means $\pm$ SD and n (%).

tions were obtained from each subject on different days, the second within the subsequent 2–28 days.

Participants ranged from 8–35 years of age, weighed at least 40 kg (88 lbs), and were free of conditions or treatments that would affect the immune system. Women were not pregnant or lactating. Control subjects with type 1 diabetes were diagnosed within the past 12 months before the first collection; control subjects did not have a first- or second-degree relative with type 1 diabetes.

Laboratory methods. The specimens were collected between 8:00–10:00 A.M. after an overnight fast, with fasting plasma glucose level between 70–180 mg/dl (3.9–10 mmol/l), and without an injection of short-acting insulin. Fresh blood samples were collected and air-shipped by overnight courier (North American sites) or same-day courier (U.K. site) at ambient temperature to each T-cell laboratory, volume permitting. Before the first visit, control participants were screened for their HLA genotypes at a central laboratory from a buccal swab. Biochemical autoantibodies (GAD-65, ICA-512, microinsulin autoantibody) were measured centrally using a radio-immunobinding assay and islet cell antibodies (ICA) using indirect immunofluorescence from frozen serum collected at the first visit (13–16). Microinsulin autoantibody was not used in this analysis because all subjects were treated with exogenous insulin. The responses in the T-cell assays were classified as positive (diagnostic of type 1 diabetes), negative (diagnostic of control status), or indeterminate.

The laboratories were masked to the status of the subject, the identity of each subject, and the sequence of visits. Detailed laboratory methods and analytes for each assay are presented in the online appendix available at http://diabetes.diabetesjournals.org/cgi/content/full/db09-0249/DC1. Herein acronyms are used to refer to specific analytes, such as "MDR4GAD274" to refer to the DR4-GAD274-286 tetramer stimulated with the DR4 binding peptide pool. The description of each analyte is presented in the online appendix, and brief methods follow.

Cellular immunoblotting. Human islets were subjected to preparative 10% SDS-PAGE, the gels were electroblotted onto nitrocellulose, nitrocellulose particles prepared, and the nitrocellulose particles used to stimulate peripheral blood mononuclear cells (PBMNCs) in vitro (7). Eighteen blot sections of decreasing size were analyzed. A stimulation index $(S_i) > 2.0$ indicated positive proliferation for a blot, and a sample with four or more positive blots was designated as positive (diagnostic of type 1 diabetes).

U.K.-ELISPOT. PBMNCs producing interferon-γ (IFN-γ) in response to stimulation with synthetic peptides representing naturally processed and presented epitopes of IA-2, GAD65, and proinsulin were detected by cytokine

ELISPOT as described (5). Only samples from HLA-DR3 or DR4+ subjects were used. Single peptides were tested at 10 μmol/l, and the pool was tested at 3 μmol/l and 10 μmol/l (final concentration of individual peptide components). A stimulation index (S_i, derived as number of spots in test analyte wells/number in negative control wells) of ≥ 3.0 was designated as a positive response to an analyte. Samples showing a response to one or more analytes were classified as positive.

T-cell proliferation assay. This assay measured antigen-induced proliferation of PBMNCs in microcultures by up to 20 individual test antigens including type 1 diabetes-relevant and -irrelevant peptides and proteins (3,8). Antigen responses were normalized as stimulation indexes, and samples with three or more T-cell pools that target type 1 diabetes-relevant antigens/epitopes (17) were designated as positive.

Tetramer. The construction of the expression vectors for generation of soluble DR0401 (DRA*0101/DRB1*0401), DR0404 (DRA1*0101/DRB1*0404), or DR0301 (DRA1*0101/DRB1*0301) molecules has been described previously (18). CD4+ T-cells from PBMNC from HLA-DR3 and/or 4+ control subjects were expanded with peptides. On day 14, the cells were stained using 10 μg/ml of PE-labeled HLA-DR0401/04 or 0301 tetramers containing a specific or a negative control peptide and then with anti–CD3-FITC and anti–CD4-PerCP antibodies. Cells were analyzed on a fluorescence-activated cell sorter calibur flow cytometer with FloJo software.

U.S.-ELISPOT. The methods described above for the U.K.-ELISPOT assay were followed. The difference in the assays included the following: *1*) Those in the U.K. were processed on the day of blood draw whereas those in the U.S.-ELISPOT were air-shipped overnight. *2*) Because of lower overall responses noted in the shipped samples compared to fresh control samples, a S_i of ≥ 2.0 was designated as a positive response to an analyte. *3*) Test analytes did not include the DR4-restricted GAD65$_{554–575}$ peptide. *4*) The test peptides were only tested at the 10 μmol/l concentration.

Statistical methods. Only evaluable specimen results were used. A sample was declared nonevaluable if the number of viable cells was inadequate or the sample was discarded because of poor quality as indicated by hemolysis before processing or dye exclusion after processing, if the yield of PBMNCs was $<0.5 \times 10^6$/ml blood, or there was no response ($S_i < 3.0$) to Pediacel (for U.K.-ELISPOT).

Standard measures of diagnostic accuracy were used (19). The assay results are either + (diagnostic of type 1 diabetes) or − for each analyte or for the completed assay. The sensitivity (proportion of patients with type 1 diabetes who test positive) and specificity (proportion of control subjects who test negative) are presented. Because it was planned that about half the
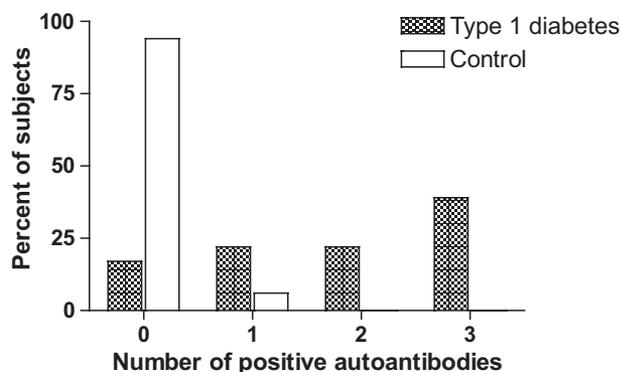
FIG. 1. The percentage of subjects with positive autoantibodies. The data show the percentage of subjects with 0–3 positive autoantibodies.

subjects with DR3/4 would have diabetes (D), the positive predictive value (PPV) = P(D +) = sensitivity/(sensitivity + α) and the negative predictive value (NPV) = P(D +) = specificity/(specificity + β). The proportion correctly classified (PCC) = (sensitivity + specificity)/2. A generalized estimating equation (GEE) logistic model (20) provided estimates of sensitivity and specificity, allowing for multiple collections for each subject, and of the odds ratio of a specimen being from a subject with diabetes given a positive classification from the assay.

The κ agreement statistic (21) assessed the reproducibility of the repeated qualitative assays (positive or negative) from the pair of collections from each control subject. A constant of 0.5 was added to all frequencies when there was a zero cell in the corresponding 2×2 table. The Cochran test of homogeneity tested the differences between the κ values among subgroups (22). An entropy R^2 measure of association (the uncertainty coefficient) described the correlation of the repeated assays (23).

For each assay, 30 subjects with type 1 diabetes and 30 or more control subjects provides 85% power with a one-sided test at the 0.05 level to detect a difference in the proportions positive of 0.675 among those with diabetes versus 0.325 (one – specificity) among those without, or an odds ratio of 4.31.

RESULTS

Study subjects. Table 1 describes the characteristics of the 94 subjects from North America and the 60 from the U.K. who contributed an evaluable specimen as defined in the online appendix. Of these, 9 in North America and 15 in U.K. withdrew after collecting the first sample. North American patients were pediatric, whereas U.K. patients were adult with somewhat lower prevalence of autoanti-

bodies. Although sensitivity for any single autoantibody ranged from 59–67%, 61% of subjects with type 1 diabetes were positive for two or more autoantibodies versus 0% of the control subjects (Table 1, Fig. 1). These observations place the test subjects well within the expected serologic ranges (15).

Discriminant validity. Table 2 presents the fraction of specimens that were evaluable for autoantibodies and each assay. Antibodies were evaluable in virtually all of the specimens. The cellular immunoblot was evaluable in 68% and T-cell proliferation (TCP) in 84% of all specimens. The other assays were only conducted in specimens from DR3 or DR4 subjects, and of these the tetramer was evaluable in 77%, U.S.-ELISPOT in 57%, and the U.K.-ELISPOT in 100%. Indeterminant specimens for which the assay did not provide a clear positive or negative response were frequent with the tetramer assay (27% of those otherwise evaluable), less so for the cellular immunoblot (5%) and U.K.-ELISPOT (7%), and none for the other two laboratories.

For the analysis herein, indeterminate specimens are counted as a negative response (see the online appendix). Table 2 presents the summary measures of diagnostic accuracy for each assay. The summary measures for each analyte within each assay are presented in the online appendix.

The cellular immunoblot assay provided sensitivity of 74%, specificity of 88%, and 81% correctly classified (Table 2), somewhat lower than provided by autoantibodies alone. Of those classified positive, 86% (the PPV) actually had type 1 diabetes, and of those classified as negative, 77% (NPV) were actually control subjects. The odds ratio (21.7) was highly statistically significant ($P < 0.0001$). Figure 2A shows the distribution of the number of positive blot sections within each group. Among control subjects, the percent with 0–3 blots positive was higher than among those with diabetes. Most of the false-positive control samples had four positive responses and were borderline for overall positivity.

The U.K.-ELISPOT assay also showed a statistically significantly odds ratio (3.44, $P = 0.0026$), with a sensitivity and specificity of 61 and 69%, respectively (Table 2).

TABLE 2

Numbers of specimens assayed and measures* of the ability to discriminate between subjects with and without type 1 diabetes for autoantibodies alone† and each T-cell assay

	Specimens evaluable‡	Indeterminant§	Sensitivity	Specificity	Correct classification	Positive predictive value	Negative predictive value	Odds ratio‖	P
Autoantibodies (one or more)	296 (99.7)	—	83	92	88	91	84	56.9	<0.0001
Anti-GAD65	296 (99.7)	—	68	98	83	98	76	124.8	<0.0001
Anti-ICA512	296 (99.7)	—	58	99	78	98	70	119.3	<0.0001
Anti-ICA	296 (99.7)	—	59	94	76	91	69	22.6	<0.0001
Cellular immunoblot	122 (68.2)	6 (4.9)	74	88	81	86	77	21.7	<0.0001
T-cell proliferation	151 (84.4)	0	60	69	64	66	63	3.36	0.0041
Tetramer	117 (76.5)	32 (27.4)	46	72	59	63	57	2.10	0.076
U.S.-ELISPOT	87 (56.8)	0	35	65	50	51	50	1.09	0.95
U.K.-ELISPOT	109 (100)	8 (7.3)	61	69	65	66	64	3.44	0.0026

Data are n (%) or percent unless otherwise indicated. *Sensitivity and specificity estimated from a GEE model using all collections in all subjects. †All subjects from both North America and the U.K. combined. ‡The number and percent of all specimens received that were considered evaluable among 154 subjects and 297 collections in North America and the U.K. combined (for autoantibodies); among 179 collections for cellular immunoblot and T-cell proliferation; and 183 DR3 and/or DR4 in North America (tetramer and U.S.-ELISPOT) and 109 in the U.K.-ELISPOT. §The number and percent of evaluable specimens for which the assay failed to provide a clearly positive or negative result. ‖The odds ratio that a specimen is from a subject with type 1 diabetes for a positive assay vs. that for a negative assay.
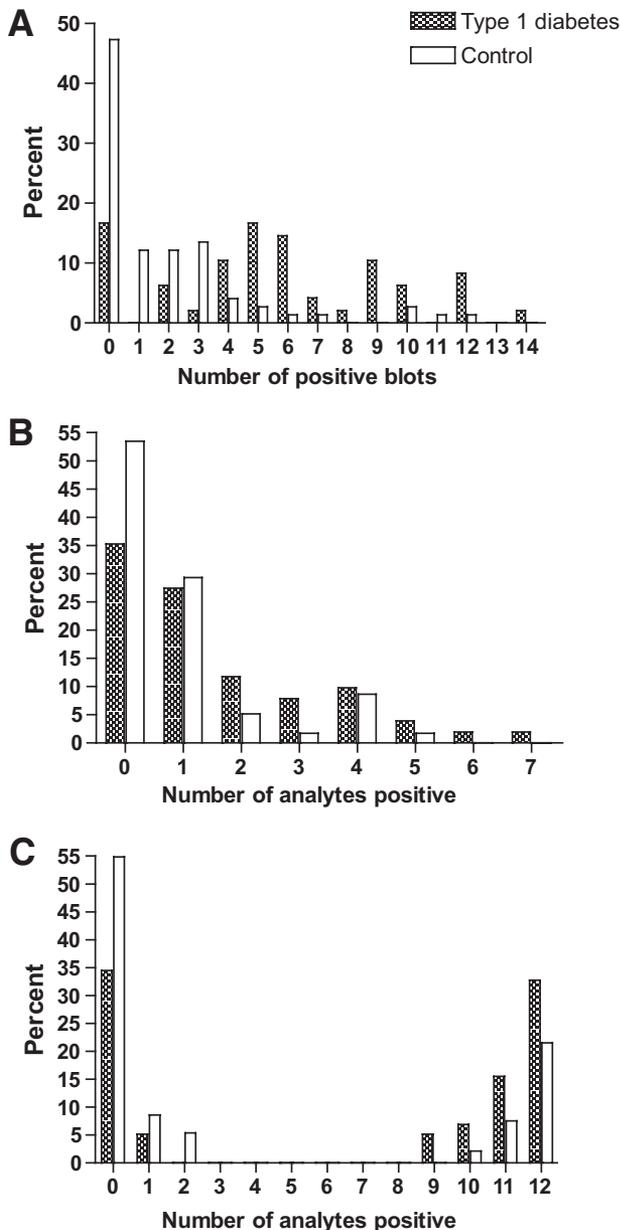
FIG. 2. The frequency of positive analytes within T-cell assays. The number of positive analytes for each of the assays are shown for patients with type 1 diabetes and healthy control subjects. *A*: Combined immunoblot assay; *B*: U.K.-ELISPOT assay; *C*: TCP assay.

Figure 2*B* shows the distribution of the number of positive responses in each group. Among the control subjects, over 80% of the collections had zero or one analyte positive. Among those with diabetes, two-thirds were positive for one or more analytes.

The TCP assay provided sensitivity of 60% and specificity of 69% and a statistically significant odds ratio (3.36, $P = 0.0041$, Table 2). Figure 2*C* shows the distribution of the number of positive responses within each group. A higher fraction of control subjects was positive for 1–3 antigens than were subjects with diabetes, whereas a higher fraction of those with diabetes was positive to 10 or more analytes than were control subjects.

The odds ratios for the tetramer (2.1) and U.S.-ELISPOT (1.09) assays were not statistically significant (Table 2). In this analysis, the sensitivity was <50% for both assays,

whereas the specificity was comparable to the other assays (Table 2).

Influence of HLA status. The major histocompatability complex genotype likely affects the performance of assays and, by design, the tetramer and ELISPOT assays. The tetramer assay relies upon HLA-restricted presentation of selected autoantigenic epitopes presented by HLA-DR3 or DR4, and the ELISPOT was optimized using epitopes presented by HLA-DR4 (DRB1*0401) with the exception of $GAD65_{335-352}$, a sequence known to be HLA-DR3 restricted (24). Therefore, we compared responses in type 1 diabetic patients and control subjects among these genotypes (Table 3). The sensitivity and specificity of each assay varied among HLA classes, but there was not a significant relationship between HLA genotype for any of the assays or autoantibodies. The cellular immunoblot assay had a higher specificity (100%) among samples from subjects with DR3 alone than those with DR4 (79% alone, 81% heterozygous). The U.K.-ELISPOT assay showed somewhat better specificity (78%) among samples from subjects with DR4 alone than subjects with DR3 (57% alone, 60% heterozygous) as previously reported (5). There was a low sensitivity (17%) but a high specificity (93%) of the tetramer assay in HLA-DR4 individuals. The small number of non-DR3 or -DR4 subjects with diabetes (only three with type 1 diabetes in North America) precluded a comparison to those with HLA-DR3 or four in the cellular immunoblot, TCP, and autoantibody assays.

Qualitative reproducibility of assays. Table 4 presents the qualitative reproducibility of the assay classifications from the repeat collections in each subject. Negative values indicate agreement less frequent than expected by chance. For example, the cellular immunoblot had 82.22% agreement for the 45 subjects with evaluable specimens on the two collections, both type 1 diabetes and normal together. The proportion positive on average for the two collections is 40%, and the corresponding level of agreement by chance alone is 51.6%. Thus, κ is (0.8222 − 0.0.516)/(1 − 0.516) = 0.633, a modest improvement over chance agreement alone but somewhat below a desirable level of 0.8 or more. The level of agreement and κ for autoantibodies was only slightly greater than that of the cellular immunoblot assay.

It is possible that these κ values represent biological variation over the interval 2–28 days between collections, rather than laboratory reproducibility. Thus, for each laboratory, separately among those with diabetes and control subjects, κ was calculated for those retested within 2–7 days, 8–14 days, and 15–28 days, and a test of homogeneity conducted. The detailed results are presented in the online appendix. Although the κ values showed nominally significant heterogeneity for four of the laboratories, there was trend toward decreasing levels of κ as the elapsed time increased only for the TCP and tetramer labs. Those for autoantibodies did not vary.

κ is not a measure of correlation between the two collections. Thus, Table 4 also presents the entropy R^2 between the two collections for each laboratory. This is a measure of the proportion of variation in values from one collection that is explained by variation in the other collection, directly analogous the square of a correlation coefficient for quantitative variables. As for κ, the R^2 is highest for the cellular immunoblot lab followed by TCP and the others.

Agreement between assays. Table 5 presents the κ and the R^2 values for agreement between laboratories. The

TABLE 3
Sensitivity and specificity of each assay among HLA classes with a test of homogeneity among classes*

	Subjects: North America/U.K.		Autoantibodies: North America and U.K. combined		Cellular immunoblot		T-cell proliferation		Tetramer		U.S.-ELISPOT		U.K.-ELISPOT	
	Diabetes	Control	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
All	35/29	59/31	83	92	74	88	60	69	46	72	35	66	61	69
DR3 alone	7/4	16/11	82	92	86	100	61	77	46	58	84	70	57	57
DR4 alone	14/16	23/17	81	92	69	79	65	73	17	93	26	61	54	78
DR3/4	11/9	9/3	87	100	60	81	56	33	74	50	26	77	72	60
Non-DR3/4	3/0	11/0	83	87	0	90	45	82	NA	NA	NA	NA	NA	NA

Data are n/n or percent. *Sensitivity and specificity estimated from a GEE model with repeat samples from each subject, separately within each HLA category. The model did not converge for the analyses of autoantibodies and the Cellular Immunoblot Laboratory for the few control subjects with non-DR3/4 owing to the absence of any positives (100% specificity). For these, the average value of sensitivity and other measures from the first and second visit are reported. NA, not applicable.

TABLE 4
The numbers of subjects with evaluable assays from the two successive visits, the simple proportion of agreement between visits, the κ index of agreement and its 95% CI, and the entropy R^2

	Agreement				Entropy
	n^*	(%)	κ	95% CI	R^2
Autoantibodies	142	86	0.70	0.58–0.82	0.40
Cellular immunoblot	45	82	0.63	0.41–0.86	0.34
U.K.-ELISPOT	49	55	0.08	−0.20 to 0.36	0.005
T-cell proliferation	63	76	0.52	0.30–0.73	0.20
Tetramer	50	74	0.41	0.15–0.68	0.13
U.S.-ELISPOT	29	66	0.25	−0.09 to 0.60	0.055

*The number of subjects for whom the laboratory considered the specimen from the two visits to be evaluable.

cellular immunoblot lab had modest κ agreement and some correlation with the presence of autoantibodies but not the other laboratories. There was weak agreement/correlation between the other labs and autoantibodies and each other, the strongest being between the cellular immunoblot and TCP laboratories but with a κ of 0.25.

We also assessed the concordance of false-positive and false-negative results among the four North American assays from split samples. Of the 113 samples from control subjects judged evaluable by at least one laboratory, 54 were assayed by three or four of the laboratories. None of these were positive in more than two of the laboratories. Likewise, among the 66 samples from those with diabetes, 47 were assayed in three or four of the laboratories, and 8 (17%) were negative in three of the labs, none in all four.
Analyte properties. The online appendix presents the discriminant validity and quantitative reproducibility of individual analytes within each laboratory assay. The quantitative discriminant ability of each analyte's numerical values were essentially the same as the analyte positivity. Figure 3*A* presents the sensitivity and specificity for each blot section in the cellular immunoblot assay. Many of the blot sections showed a specificity greater than 80%, some greater than 90%, but the maximum sensitivity was 50% and the maximum reliability coefficient was 60% (both for blot 16). For the U.K.-ELISPOT assay (Fig. 3*B*), no analytes had a sensitivity >30%, whereas all the analytes had a specificity >80%. The reproducibility of responses to analytes was also limited, the highest being 21%. For the TCP assay (Fig. 3*C*), the maximum sensitivity, specificity, and reliability of any analytes were 64, 75, and 42%, respectively, all for different peptides or protein antigens.

For the tetramer assay, the maximum sensitivity was 89% but in an assay with low specificity (7%). The maximum specificity for one analyte was 100% with 32% sensitivity, both greater per individual analyte than any of the other HLA-dependent assays. The maximum reliability was 74% but in an assay with lower sensitivity (50%). For the U.S.-ELISPOT assay, all of the analytes showed specificity of at least 77%, but the maximum sensitivity was 18% and the maximal reliability 44%.

DISCUSSION

The reliable detection of T-cell responses associated with the autoimmune processes that lead to type 1 diabetes has been a major research goal for nearly 2 decades. In a blinded controlled study, we tested the ability of five T-cell

TABLE 5
Indexes of agreement between each pair of laboratories in North America for the designation of positive vs. negative for each subject based on the first visit for which each pair of labs both provided a result

	Autoantibodies	Cellular immunoblot	TCP	Tetramer	U.S.-ELISPOT
Autoantibodies	—	0.51 (64)	0.10 (81)	0.11 (59)	0.11 (45)
Cellular immunoblot	0.20	—	0.25 (67)	0.14 (53)	−0.012 (46)
TCP	0.008	0.05	—	0.09 (63)	0.03 (55)
Tetramer	0.009	0.015	0.007	—	−0.29 (47)
U.S.-ELISPOT	0.009	0.0001	0.001	0.07	—

Data are κ (n of concurrent evaluations) and entropy R^2. $\kappa < 0$ designates worse-than-chance agreement.

assays to distinguish responses in participants with type 1 diabetes from normal healthy control subjects. We also assessed the qualitative reproducibility of the overall measurement of positivity and the quantitative reproducibility of measured responses to individual assay analytes to understand the nature of the responses to diabetes antigens. The CI, TCP, and U.K.-ELISPOT assays could distinguish between persons with type 1 diabetes from control subjects with sensitivity and specificity ranging between 60–74% and 69–88%, respectively. We did not find a significant effect of HLA genotypes on responses in the combined immunoblot and TCP assays. The ELISPOT assay utilized DR4-restricted peptides predominantly and showed a trend for greater specificity among DR4+ individuals. The tetramer assay was less discriminatory but showed high specificity in DR4+ individuals, possibly reflecting the selection of peptides for binding to this genotype.

Several explanations could account for the differences between the ELISPOT assay results in the U.K. and Denver, Colorado. The U.K. subjects were older and had shorter disease duration. Samples were transported to the U.K. lab on the day they were drawn, whereas the Denver samples were air-shipped, exposed to temperature variation, and more laboratory personnel were involved. These technical details highlight the challenges in replicating bioassays at different laboratory sites.

The performance characteristics with these biologic assays are similar to other biologic assays and interestingly not markedly different from the biochemical autoantibody assays (25,26). Of note, the study design tested the reproducibility of assay results from the same subjects and not the reproducibility of measurements in the same samples. In this regard, the qualitative and quantitative reproducibility of assays in individual subjects between the two samplings and the individual analyte reproducibility were lower than desirable, with the maximum κ agreement statistic of 0.63 and maximum R^2 of 0.34, both for the cellular immunoblot assay. The reasons for this variation are not clear; they could reflect variation in the assay or in procedures used to ship or process samples, but they could also reflect biological variation in individuals over time. This is further suggested by 0.7 κ for the autoantibodies, whereas in split duplicates, the κ values were 0.89 and 0.93 for anti-GAD65 and ICA512. Nonetheless, a systematic decline in the κ statistics as a function of elapsed duration was only seen in two laboratories. Thus, if biological mechanisms explain the variation, they must take place within days.

In a previous study, Seyfert-Margolis et al. suggested that "false"-positive results in cellular assays may not be false but a true measure of diabetes antigen–specific T-cells in normal individuals because 2 of 4 positive samples from normal control subjects in the combined immunoblot were also positive in the TCP assay (8). However, in the present study, the low concordance among laboratories in false-positive responses for control subjects suggests that false positives are assay specific rather than a true biologic difference in these healthy individuals. Likewise, the low concordance of false negatives between subjects suggests that they are assay specific.

Each of the assays measures the T-cell responses in ways that differ with respect to assay conditions and antigenic complexity. The combined immunoblot measures the proliferative responses of peripheral blood cells to all islet antigens, whereas the ELISPOT, TCP, and tetramer assays measure responses to a limited repertoire; only a fraction of the cells may be present in the peripheral blood at any time (3–5,7,27,28). The responding cell subpopulations (CD4+, CD8+, or both) also differ between the assays. Therefore, it is not surprising that there was not complete agreement between results with each assay. The relatively higher sensitivity of the combined immunoblot compared to either the T-cell proliferative, ELISPOT, or tetramer assays most likely reflects their detection of responses to the widest array of antigens. A caveat of these studies is that we have not compared responses in patients with type 1 diabetes with other autoimmune diseases. In this exercise, therefore, we cannot distinguish diabetes-specific responses from an "autoimmune" phenotype.

There are limitations of these assays. First, although the highest sensitivity and specificity was with the combined immunoblot assay, it is not clear which antigens are recognized. The combined immunoblot and TCP assays identify the number of positive analytes rather than the responses to each analyte. In this regard, the ELISPOT and tetramer assays may be more useful for following the responses of particular antigen-specific cells. Moreover, the limited reproducibility of the combined immunoblot and TCP assays to specific analytes, whether this is because of biological shifts in particular T-cell populations or technical variability, may present a problem for tracking responses to any group of antigen-specific T-cells over the course of a clinical study. On a practical level, the volume of blood needed for each of the assays (ranging between 10–30 ml) may present a problem for repeated sampling on smaller subjects.

These assays all used fresh rather than frozen PBMNC. In a previous analysis, assays that used frozen cells did not perform well, possibly because of variability in cryopreservation procedures, whereas both the combined immunoblot and TCP assays that used fresh cells showed good discriminant ability, particularly when combined (8). Be-
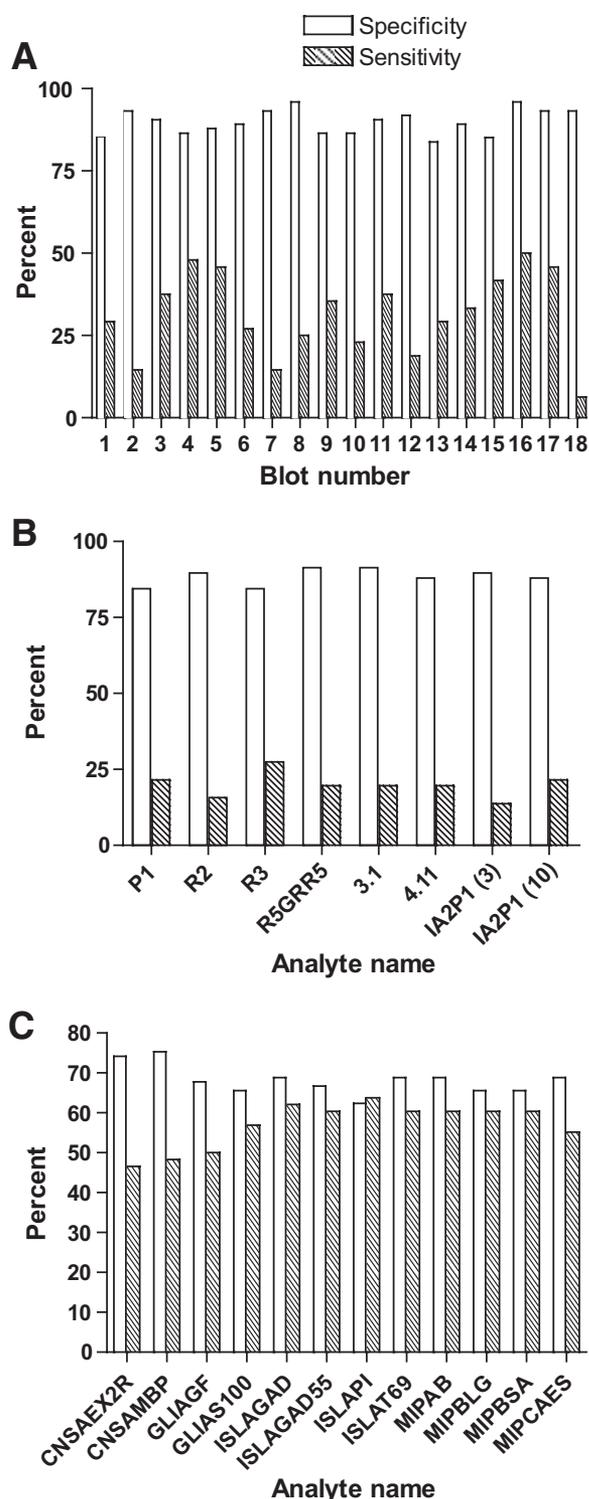
FIG. 3. Sensitivity and specificity of each of the analytes used in the T-cell assays. *A*: Combined immunoblot assay; *B*: U.K.-ELISPOT assay; *C*: TCP assay.

cause using fresh cells limits performance of the assay to the time of the sampling, the reproducibility of the assay over time, particularly of the individual analytes, could complicate the use of these assays in clinical trials. An analysis of the many thousands of (strictly blinded) TCP samples processed over years in the TRIGR diabetes prevention trial may provide new answers to these questions (29).

A number of improvements might be made in the future to enhance the use of these assays in clinical studies. Clearly optimizing methods for freezing cells and adapting the assays for use with frozen cells is important since it would allow the simultaneous measurement of samples from individual subjects collected at different points in time. With selection of the most informative analytes, the volumes of blood needed to run the assays could be reduced. Importantly, repeated studies over time would help to understand the appearance and disappearance of antigen-specific T-cells in the peripheral circulation and help to understand whether the changes that were seen in the assay results reflect differences in the handling of the specimens or reflect a biologic change that may be because of trafficking of antigen-specific cells through various compartments. Nonetheless, our findings provide highly encouraging results regarding the ability of cellular assays to identify responses to multiple targets that can discriminate patients from normal control subjects. Although the discriminatory ability is still superior with a combination of autoantibody measurements, the cellular assays provide insights into cells that are thought to be involved in the disease pathogenesis and are likely to be affected by new interventions that target immune responses.

REFERENCES

1. Atkinson MA. ADA Outstanding Scientific Achievement Lecture 2004. Thirty years of investigating the autoimmune basis for type 1 diabetes: why can't we prevent or reverse this disease? Diabetes 2005;54:1253–1263
2. Ellis T, Jodoin E, Ottendorfer E, Salisbury P, She JX, Schatz D, Atkinson MA. Cellular immune responses against proinsulin: no evidence for enhanced reactivity in individuals with IDDM. Diabetes 1999;48:299–303
3. Dosch H, Cheung RK, Karges W, Pietropaolo M, Becker DJ. Persistent T cell anergy in human type 1 diabetes. J Immunol 1999;163:6933–6940
4. Reijonen H, Novak EJ, Kochik S, Heninger A, Liu AW, Kwok WW, Nepom GT. Detection of GAD65-specific T-cells by major histocompatibility complex class II tetramers in type 1 diabetic patients and at-risk subjects. Diabetes 2002;51:1375–1382
5. Arif S, Tree TI, Astill TP, Tremble JM, Bishop AJ, Dayan CM, Roep BO, Peakman M. Autoreactive T cell responses show proinflammatory polarization in diabetes but a regulatory phenotype in health. J Clin Invest 2004;113:451–463
6. Alleva DG, Crowe PD, Jin L, Kwok WW, Ling N, Gottschalk M, Conlon PJ, Gottlieb PA, Putnam AL, Gaur A. A disease-associated cellular immune response in type 1 diabetics to an immunodominant epitope of insulin. J Clin Invest 2001;107:173–180
7. Brooks-Worrell BM, Starkebaum GA, Greenbaum C, Palmer JP. Peripheral blood mononuclear cells of insulin-dependent diabetic patients respond to multiple islet cell proteins. J Immunol 1996;157:5668–5674
8. Seyfert-Margolis V, Gisler TD, Asare AL, Wang RS, Dosch HM, Brooks-Worrell B, Eisenbarth GS, Palmer JP, Greenbaum CJ, Gitelman SE, Nepom GT, Bluestone JA, Herold KC. Analysis of T-cell assays to measure autoimmune responses in subjects with type 1 diabetes: results of a blinded controlled study. Diabetes 2006;55:2588–2594
9. Kaufman DL, Tisch R, Sarvetnick N, Chatenoud L, Harrison LC, Haskins K, Quinn A, Sercarz E, Singh B, von Herrath M, Wegmann D, Wen L, Zekzer D. Report from the 1st International NOD Mouse T-Cell Workshop and the follow-up mini-workshop. Diabetes 2001;50:2459–2463
10. Schloot NC, Meierhoff G, Karlsson Faresjo M, Ott P, Putnam A, Lehmann P, Gottlieb P, Roep BO, Peakman M, Tree T. Comparison of cytokine ELISpot assay formats for the detection of islet antigen autoreactive T-cells. Report of the third immunology of diabetes society T-cell workshop. J Autoimmun 2003;21:365–376

11. Staeva-Vieira T, Peakman M, von Herrath M. Translational mini-review series on type 1 diabetes: immune-based therapeutic approaches for type 1 diabetes. Clin Exp Immunol 2007;148:17–31

12. Lambert AP, Gillespie KM, Thomson G, Cordell HJ, Todd JA, Gale EA, Bingley PJ. Absolute risk of childhood-onset type 1 diabetes defined by human leukocyte antigen class II genotype: a population-based study in the United Kingdom. J Clin Endocrinol Metab 2004;89:4037–4043

13. Yu L, Rewers M, Gianani R, Kawasaki E, Zhang Y, Verge C, Chase P, Klingensmith G, Erlich H, Norris J, Eisenbarth GS. Antiislet autoantibodies usually develop sequentially rather than simultaneously. J Clin Endocrinol Metab 1996;81:4264–4267

14. Verge CF, Gianani R, Kawasaki E, Yu L, Pietropaolo M, Jackson RA, Chase HP, Eisenbarth GS. Prediction of type I diabetes in first-degree relatives using a combination of insulin, GAD, and ICA512bdc/IA-2 autoantibodies. Diabetes 1996;45:926–933

15. Eisenbarth GS, Gianani R, Yu L, Pietropaolo M, Verge CF, Chase HP, Redondo MJ, Colman P, Harrison L, Jackson R. Dual-parameter model for prediction of type I diabetes mellitus. Proc Assoc Am Physicians 1998;110:126–135

16. Krischer JP, Cuthbertson DD, Yu L, Orban T, Maclaren N, Jackson R, Winter WE, Schatz DA, Palmer JP, Eisenbarth GS. Screening strategies for the identification of multiple antibody-positive relatives of individuals with type 1 diabetes. J Clin Endocrinol Metab 2003;88:103–108

17. Winer S, Tsui H, Lau A, Song A, Li X, Cheung RK, Sampson A, Afifiyan F, Elford A, Jackowski G, Becker DJ, Santamaria P, Ohashi P, Dosch HM. Autoimmune islet destruction in spontaneous type 1 diabetes is not β-cell exclusive. Nat Med 2003;9:198–205

18. Novak EJ, Liu AW, Nepom GT, Kwok WW. MHC class II tetramers identify peptide-specific human CD4(+) T cells proliferating in response to influenza A antigen. J Clin Invest 1999;104:R63–R67

19. Pepe MS. *The Statistical Evaluation of Medical Tests for Classification and Prediction.* Oxford, UK, Oxford University Press, 2003

20. Diggle PJ, Heagerty PJ, Liang K-Y, Zeger SL. *Analysis of Longitudinal Data.* Oxford, UK, Oxford University Press, 2002

21. Fleiss JL, Levin B, Paik MG. *Statistical Methods for Rates and Proportions.* New York, NY, John Wiley and Sons, 2003

22. Lachin JM. *Biostatistical Methods: The Assessment of Relative Risks.* New York, NY, John Wiley and Sons, 2000

23. Lachin JM. The role of measurement reliability in clinical trials. Clinical Trials 2004;1:553–566

24. Di Lorenzo TP, Peakman M, Roep BO. Translational mini-review series on type 1 diabetes: systematic analysis of T cell epitopes in autoimmune diabetes. Clin Exp Immunol 2007;148:1–16

25. Fritzler MJ, Wiik A, Tan EM, Smolen JS, McDougal JS, Chan EK, Gordon TP, Hardin JA, Kalden JR, Lahita RG, Maini RN, Reeves WH, Rothfield NF, Takasaki Y, Wilson M, Byrd MG, Slivka L, Koziol JA. A critical evaluation of enzyme immunoassay kits for detection of antinuclear autoantibodies of defined specificities. III. Comparative performance characteristics of academic and manufacturers' laboratories. J Rheumatol 2003;30:2374–2381

26. Lopez-Hoyos M, Rodriguez-Valverde V, Martinez-Taboada V. Performance of antinuclear antibody connective tissue disease screen. Ann N Y Acad Sci 2007;1109:322–329

27. Brooks-Worrell B, Gersuk VH, Greenbaum C, Palmer JP. Intermolecular antigen spreading occurs during the preclinical period of human type 1 diabetes. J Immunol 2001;166:5265–5270

28. Reijonen H, Kwok WW. Use of HLA class II tetramers in tracking antigen-specific T cells and mapping T-cell epitopes. Methods 2003;29:282–288

29. Akerblom HK, Virtanen SM, Ilonen J, Savilahti E, Vaarala O, Reunanen A, Teramo K, Hamalainen AM, Paronen J, Riikjarv MA, Ormisson A, Ludvigsson J, Dosch HM, Hakulinen T, Knip M. Dietary manipulation of β cell autoimmunity in infants at increased risk of type 1 diabetes: a pilot study. Diabetologia 2005;48:829–837