
Peer reviewed version

Link to published version (if available):
10.1373/clinchem.2019.304196

Link to publication record in Explore Bristol Research

PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via American Association for Clinical Chemistry at http://clinchem.aaccjnls.org/content/65/9/1141. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/user-guides/explore-bristol-research/ebr-terms/
Islet Autoantibody Standardization Program 2018 workshop: interlaboratory comparison of glutamic acid decarboxylase autoantibody assay performance

Lampasona V\textsuperscript{a\,*}, Pittman DL\textsuperscript{b}, Williams AJ\textsuperscript{c}, Achenbach P\textsuperscript{d}, Schlosser M\textsuperscript{ef}, Akolkar B\textsuperscript{g}, Winter WE\textsuperscript{b}, and participating laboratories

\textsuperscript{a}San Raffaele Diabetes Research Institute, San Raffaele Scientific Institute, Milan, Italy; \textsuperscript{b}Department of Pathology, University of Florida, Gainesville, FL, USA, \textsuperscript{c}Diabetes and Metabolism, Translational Health Sciences, University of Bristol, UK; \textsuperscript{d}Institute of Diabetes Research, Helmholtz Center Munich, German Research Center for Environmental Health, Neuherberg, Germany; \textsuperscript{e}Department of Surgery, University Medical Center Greifswald, Greifswald, Germany; \textsuperscript{f}Institute of Pathophysiology, Research Group of Predictive Diagnostics, University of Greifswald, Karlsburg, Germany; \textsuperscript{g}Division of Diabetes, Endocrinology, and Metabolic, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, USA

\textsuperscript{\ast}These authors contributed equally

\textsuperscript{*}Address correspondence to this author: Vito Lampasona, San Raffaele Diabetes Research Institute, via Olgettina 60, 20132 Milano, Italy, telephone: +39 02 2643 2540, fax: +39 02 2643 4351, e-mail: lampasona.vito@hsr.it
Running Title: IASP2018 GAD65 autoantibody interlaboratory comparison study

Keywords: GADA, Autoantibodies, Sensitivity, Specificity, IASP Interlaboratory comparison study, type 1 diabetes

Abbreviations: AC1, Gwet’s coefficient of inter-rater agreement reliability; APPA, average pairwise percent agreement; GADA, autoantibodies to glutamic acid decarboxylase GAD65 protein; IASP, Islet Autoantibody Standardization Program; ROC-AUC, Area Under the ROC Curve; AS95, Adjusted Sensitivity at 95% specificity; IQR, interquartile range; k, Fleiss’ coefficient of inter-rater agreement reliability; OCCC, Overall Concordance Correlation Coefficient; W, Kendall’s ranking agreement coefficient; RBA, Radio Binding Assay; ECL, Electro Chemi Luminescence; LIPS, Luciferase Immuno Precipitation System; MPNIRF, Multiplex Plasmonic Near-InfraRed Fluorescence; CLIA, Chemi Luminescence Immuno Assay; LBI, Luminex Bead Immunoassay; MFRET, Multiplex Fluorescence Energy Transfer
Abstract

Background
The islet autoantibody standardization program (IASP) aims to improve the performance of immunoassays measuring type 1 diabetes (T1D) associated autoantibodies and the concordance of results among laboratories. IASP organizes international interlaboratory assay comparison studies in which blinded serum samples are distributed to participating laboratories, followed by centralized collection and analysis of results, providing participants with an unbiased comparative assessment. In this report we describe the results of glutamic acid decarboxylase autoantibodies (GADA) assays presented in the IASP 2018 workshop.

Methods
In May 2018, IASP distributed to participants uniquely coded sera from 43 new onset T1D patients, 7 multiple autoantibody positive non-diabetic subjects, and 90 blood donors. Results were analyzed for the following metrics: sensitivity, specificity, accuracy, area under the receiver-operator characteristic curve (ROC-AUC), partial ROC-AUC at 95% specificity (pAUC95), concordance of qualitative and quantitative results.

Results
37 laboratories submitted results from a total of 48 different GADA assays adopting 9 different formats. The median ROC-AUC and pAUC95 of all assays were 0.87 (interquartile range 0.83 – 0.89) and 0.036 (interquartile range 0.032 – 0.039), respectively. Large differences in pAUC95 (range 0.001 – 0.0411) were observed across assays. Of formats widely adopted, bridge-ELISAs showed the best median pAUC95 (0.039, range 0.036 – 0.041).

Conclusions
Several novel assay formats submitted to this study showed heterogeneous performance. In 2018 the majority of the best performing GADA immunoassays consisted of novel or
established non-radioactive tests that proved on a par or superior to the Radio Binding Assay,

the previous gold standard assay format for GADA measurement.
Introduction

The Islet Autoantibody Standardization Program (IASP) is a collaborative effort aimed at improving the performance of assays measuring Type 1 diabetes (T1D) associated autoantibodies and the concordance of results between laboratories (1).

IASP is supported by the Immunology of Diabetes Society (IDS) and the NIH, coordinated by an IDS nominated committee and run by the University of Florida Pathology Laboratories, Endocrine Autoantibody Laboratory.

IASP organizes international interlaboratory comparison studies in which blinded T1D and control serum samples are tested for T1D associated autoantibodies by participating laboratories. Centralized collection and analysis of results by the IASP committee provide participants with an unbiased comparison of assay performance. Moreover, IASP fosters the continuous improvement of T1D autoantibody immunoassays through the dissemination of empirically tested best practice protocols, state of the art reagents, and serum standards.

In this report we analyze the results of assays for antibodies to glutamic acid decarboxylase 65 (GADA)(2) submitted in 2018 to the IASP interlaboratory comparison study and presented at the IASP 2018 workshop held at the 16th Immunology of Diabetes Society Congress in London, UK.

GADA are found in several neurological and endocrine autoimmune diseases (3–5). In the setting of autoimmune diabetes, GADA are the most prevalent autoantibody at onset of T1D and the hallmark of latent autoimmune diabetes in adults (LADA)(6), a slowly progressing form of pancreatic endocrine autoimmunity affecting up to 5% of type 2 diabetes patients. Moreover, GADA measurement is a cornerstone of screening strategies for T1D (7).
The most recent IASP GADA interlaboratory comparison and standardization study took place in 2018. 37 laboratories from 17 countries in North America, Europe, Asia and Australia submitted results from 48 different GADA assays, based on 9 different assay formats, after testing blinded samples from 50 cases with T1D or multiple islet autoimmunity and 90 blood donors.
Materials and methods

Study design:

In the 2018 IASP interlaboratory comparison study, participants received sets of the same serum samples consisting of: 50 cases (43 sera from new onset T1D patients and 7 multiple islet autoantibody positive first-degree relatives of T1D patients enrolled in the TrialNet Ancillary Study – Pathway to Prevention, who during screening showed a transiently altered glucose tolerance test), 90 control samples (all blood donors), and 10 additional samples to be used for sub-studies unrelated to GADA testing.

The T1D patients had a median age of 14 years (range 8-47) and included 15 females and 28 males, of whom 37 were white, 2 black, 2 of mixed and 2 of undisclosed ancestry. The multiple T1D autoantibody positive subjects had a median age of 16 years (range 12–53) and included 4 females and 3 males, all of white ancestry. The blood donors had a median age of 20 years (range 18–30) and included 44 females and 44 males, of whom 69 were white, 19 black, and 2 for whom demographic data were not available.

New onset T1D samples were contributed by several centers around the world and collected within 14 days of starting insulin treatment. Blood donor samples were collected in the USA and included only people without diabetes. All serum samples submitted to the IASP repository were collected upon obtaining a written informed consent and with the approval of local ethics committees as required by local regulations according to the ethical principles for medical research involving human subjects of the Declaration of Helsinki.

All sera were labeled and distributed as blinded 105 μL frozen aliquots, labelled with an aliquot specific unique code. Laboratories were free to use any GADA assay format but were asked to
provide details of their assay protocol and to report assay results, including raw data, to IASP for analysis using uniform Excel reporting sheets.

**Data analysis:**

We calculated sensitivity and specificity for each assay as the percentage of case sera reported as GADA-positive and as the percentage of blood donor sera reported as negative, respectively. Adjusted sensitivity 95 (AS95), i.e. the level of sensitivity corresponding to a specificity of 95%, was derived by placing the threshold for positivity at the 95th percentile of values observed in the blood donor samples in each assay.

Concordance of laboratory-assigned positive or negative scores across assays was expressed as average pairwise percent agreement between assays (APPA), i.e. the average number of times each possible combination of two assays agreed on GADA positive/negative scores divided by the number of samples scored. We tested the occurrence of agreement by pure chance by calculating the AC1 coefficient according to Gwet (8,9) and the k coefficient according to Fleiss (10) using the corresponding functions (http://agreestat.com/r_functions.html, Advanced Analytics LLC, MD, USA) in the R language and environment for statistical computing and graphics (11).

Assay performance in discriminating health from disease was analyzed using the area under the receiver-operating characteristic curve (ROC-AUC) and the partial ROC-AUC imposing a specificity greater than 95% (pAUC95)(12).

The analysis of inter-assay antibody titer concordance was performed after ranking of patient and control samples according to autoantibody levels in each assay by calculating the Kendall W rank correlation coefficient (13) using the vegan R package (https://CRAN.R-project.org/package=vegan). The significance of differences in mean ranking of selected cases
between hybrid solid/liquid phase vs liquid phase only assays was tested using the Mann Whitney test. This synthetic index based on ranking was preferred to classical pairwise regression analysis in light of the numerous comparisons to be made and to the presence of systematic differences in measurement of high and low GADA titers across assays, mostly related to the adoption of a variety of alternative algorithms for calculating local arbitrary units.

For laboratories that reported results in WHO units (14), the concordance of antibody titers was evaluated by calculating the overall concordance correlation coefficient (OCCC) according to Barnhart (15) using the f.analysis macro (16). The OCCC measures how far the fitted linear relationship of two variables X and Y deviates from the concordance line (accuracy) and how far each observation deviates from the fitted line (precision) and ranges from 0 to ±1, where results close to ±1 stand for near perfect concordance (or perfect discordance) while 0 stands for no correlation.

For all statistical analyses, two-tailed $P$-values <0.05 were considered as significant.
Results:

Summary of submitted GADA assay formats

41 laboratories from 17 countries in North America, Europe, Asia and Australia registered for the IASP2018 interlaboratory comparison study. Of the participating laboratories, 37 submitted results (online Supplemental Appendix 1) from a total of 48 different GADA assays using 9 different formats.

The submitted assays adopted the following formats: Radio Binding Assay (RBA) (18 assays, 37.5%)(17); bridge-ELISA (12 assays, 25.0%)(18); Luciferase Immuno Precipitation System (LIPS) (8 assays, 16.7%)(19); Electro Chemi Luminescence (ECL) (3 assays, 6.3%, of which 1 multiplexed individual GADA measurement of the IgG, IgA, IgM immunoglobulin classes)(20); Chemi Luminescence Immuno Assay (CLIA) (2 assays, 4.2%, of which 1 used a bridge format); Luminex Bead Immunoassay (LBI) (2 assay, 4.2%); Multiplex Plasmonic Near-Infrared Fluorescence (MPNIRF) (1 assay, 2.1%)(21); Multiplex Fluorescence Energy Transfer (MFRET) (1 assay, 2.1%); Antibody Dependent Agglutination PCR (ADAP) (1 assay, 2.1%) (22). Antigen-antibody binding occurred in liquid phase in 29 assays (ADAP, ECL, LIPS, RBA) followed by the capture of immune-complexes either through the recovery of immunoglobulins (LIPS, RBA) or tagged antigen (ECL), in hybrid solid/liquid phase in 13 assays (bridge-ELISA and CLIA), solid phase in 3 assays (LBI and MPNIRF), while binding phase was not specified for 1 assay (MFRET). Major characteristics and metrics of each individual assay are reported in online Supplemental Table 1.

Among RBAs, 5 assays adopted the NIDDK harmonized protocol (23) and 4 used truncated GAD65 antigens (24–26). Most RBAs used GAD antigens transcribed and translated in vitro (17/18 assays) and radiolabeled with $^{35}$S or $^{3}$H (16 and in 1 cases, respectively). One assay used $^{125}$I iodinated recombinant GAD65 antigen. Results in WHO units (14) were reported for all
bridge-ELISA assays, 5 RBAs and 3 LIPS that used DK standards (23). Among LIPS, 7 assays used the Nanoluc and 1 the Renilla luciferase reporters. Most LIPS assays used truncated GAD65 antigens corresponding to GAD amino acids 96-585 (GAD96-585, 4 assays) or 188-585 (GAD188-585, 2 assays) instead of full length GAD65 (GAD1-585, 1 assay).

Assay sensitivity, specificity, and accuracy

The median laboratory assigned assay sensitivity, specificity and accuracy were 69% (interquartile range 64 – 76%), 98.9% (interquartile range 96.7 – 100), and 88.6% (interquartile range 84.5 – 90.7), respectively (Table 1, Figure 1, and online Supplemental Table 1).

Assay performance

The median area under the ROC curve (ROC-AUC) was 0.87 (interquartile range 0.83 – 0.89) (Table 1, online Supplemental Table 1 and Supplemental Figure 1). As a more relevant proxy of assay performance based around commonly adopted thresholds for positivity, we calculated also the partial ROC-AUC after imposing a specificity ≥ 95% (pAUC95). The median of pAUC95 GADA assays was 0.036 (interquartile range 0.032 – 0.039) against a theoretical maximum of 0.05, and the stratification according to format highlighted a wide heterogeneity of performance across both assays and formats when high specificity was levied (Figure 2 and online Supplemental Figure 2).

The median pAUC95 of RBA assays was 0.0349 (range 0.0253 – 0.0394). The RBAs with pAUC95 above the overall median included assays using a truncated GAD65 antigen corresponding to amino acids 96-585 (4/4 assays), RBAs using the NIDDK harmonized protocol (2/5 assays), and only two RBAs using local protocols and full length GAD65 (2/8 assays).
Bridge-ELISA assays pAUC95 showed a median value of 0.0393 (range 0.0358 – 0.0411), always at or above the median of all assays.

ECL assays that, like the bridge-ELISA, are theoretically capable of detecting autoantibodies of any immunoglobulin class showed a pAUC95 of 0.0362 (range 0.0347 – 0.0377). A third ECL assay aimed to multiplex the measurement of specific GADA immunoglobulin classes (IgG, IgM, and IgA) and showed variable pAUC95 (ECL-IgG pAUC95 = 0.0304, ECL-IgM pAUC95 = 0.0160, ECL-IgA pAUC95 = 0.0120).

LIPS assays using a truncated antigen corresponding to GAD65 amino acids 96-585 and a nanoluciferase reporter showed a median pAUC95 of 0.0360 (range 0.0359 – 0.0368) while the LIPS using a Renilla luciferase reporter showed lower performance (pAUC95 = 0.0196). A LIPS assay using full-length GAD65 showed a pAUC95 of 0.034 while two LIPS assays using a truncated antigen corresponding to GAD65 amino acids 188-585 had a median pAUC95 of 0.034 (range 0.034 – 0.035).

Among assays that used formats submitted for the first time to the IASP interlaboratory comparison study, the best performance was achieved by the ADAP assay (pAUC95 = 0.0411), followed by two assays using the CLIA format with a median pAUC95 of 0.035 (range 0.031 – 0.038), while a poorer performance was observed for the three remaining assays adopting the MPNIRF (pAUC95 = 0.014), LBI (pAUC95 = 0.006), and MFRET (pAUC95 = 0.001) formats, respectively.

**Concordance of laboratory-assigned positive/negative scores**

In cases, the average pairwise percent agreement (APPA) of positive/negative scores across all assays was 83.16% and the first-order agreement coefficient AC1 (9) was 0.69. Concordance increased when the analysis was limited to assays with pAUC95 above the median (APPA =
90.82%; AC1 = 0.85) or when assays using the same format and/or antigens were compared (APPA range = 84.0 - 96.31%; AC1 range = 0.73 - 0.94) (online Supplemental Table 2).

In control samples, the APPA and AC1 coefficient across all assays were 96.10% and 0.96, respectively. Similar to cases, both agreement measures increased when assays with pAUC95 above the median (APPA = 97.52%; AC1 = 0.97) or assays using the same format and/or antigens were compared (APPA range = 93.3 - 98.2%, AC1 range = 0.92 - 0.98) with the exception of ECL and CLIA assays (online Supplemental Table 3).

The results of the Fleiss’ κ concordance coefficient (10), an alternative metric of agreement, were invariably lower (range -0.03 to 0.49), an observation consistent with the majority of control samples being scored GADA positive sporadically in only a small fraction of assays (27).

The analysis of laboratory assigned positive/negative scores highlighted the existence of format associated patterns. A subset of cases was recognized as GADA positive predominantly by ADAP and bridge-ELISA assays (samples IDS324, IDS309, IDS290, IDS312, IDS337) and to a lower extent by some ECL (samples IDS309, IDS290, IDS312, IDS337) or LIPS assays (samples IDS312, IDS337) but not by the majority of RBAs (Figure 3). Conversely, a subset of controls was recognized as GADA positive exclusively by a minority of RBAs (samples TS23727, N59416, N53371) or LIPS assays (sample S8650) (Figure 4).

To evaluate the impact of local threshold selection criteria on assay sensitivity and specificity, the data were re-analyzed after imposing a 95% specificity onto IASP2018 controls. The newly assigned positive/negative scores highlighted an improved concordance between bridge-ELISAs or ECLs and RBAs or LIPS, with all samples previously recognized as positives exclusively by the hybrid solid/liquid phase bridge-ELISA format scoring positive also in liquid phase assays (online Supplemental Figure 3). In control samples, the use of these novel
thresholds made more apparent the presence of weak format and/or antigen specific reactivities that lead to frequent positive scores for subsets of control samples in local or harmonized protocol RBAs (samples TS23727, N59416, N53371) and RBA or LIPS assays using truncated GAD96-585 (samples LQ23340, N59534, S8650) (online Supplemental figure 4).

Concordance of autoantibody titer ranks:

The inter-assay concordance of antibody titer was evaluated by first ranking sera in each assay (Figure 5A and online Supplemental Figures 5-6) followed by calculation of the Kendall’s W ranking agreement coefficient (13). The W coefficients across all assays were 0.80 and 0.13 in cases and control samples, respectively (online Supplementary Tables 2-3).

The exclusion from the analysis of assay formats with the four lowest pAUC95 led to a modest increase of the agreement coefficient in both cases and controls (W: 0.85 and 0.14, respectively) while limiting the analysis to assays with performance above the overall median pAUC95, showed again an increase of W in cases (W = 0.93) but only a marginal improvement in controls (W = 0.15), suggesting that higher concordance in cases was partially correlated with assay performance.

Concordance of GADA titer ranks increased among assays using the same format both in cases and, at least for assays other than the bridge-ELISA, also in control samples (online Supplemental Table 2-3 and online Supplemental Figures 5-6).

A marginal increase of W in cases after excluding assays measuring only IgM or IgA antibodies (W = 0.82) was observed, indicating that discrepancies of GADA titer ranks across assays were not exclusively associated with measurement of immunoglobulins of a class other than IgG.

Stratification of assays according to antibody/antigen binding in hybrid solid/liquid phase, i.e. bridge-ELISAs, or liquid phase, i.e. including ECL, RBA, and LIPS, highlighted a significant
difference between hybrid solid/liquid or liquid phase assays in mean GADA titer ranks assigned to a subset of T1D sera (IDS285, IDS301, IDS319, IDS325; Mann Whitney test: all \( P < 0.0001 \)).

**Concordance of autoantibody units:**

The concordance of arbitrary units assigned to cases was assessed across assays expressing results in WHO units. These included bridge-ELISAs, all commercial assays using standards calibrated against the NIBSC 97/550 WHO reference serum (14), and a subset of RBA and LIPS using the NIDDK DK standards, also calibrated against the WHO reference serum. Across all assays, inter-assay concordance of WHO units was relatively low with an overall correlation concordance coefficient (OCCC) of 0.4870.

After assay stratification according to antibody/antigen binding in hybrid solid/liquid phase or liquid phase, concordance of antibody titers increased across liquid phase assays adopting the NIDDK calibrators (OCCC = 0.82) compared to bridge-ELISAs (OCCC = 0.45) (Figure 5B).

Stratification of assays confirmed a significant difference between hybrid solid/liquid or liquid phase assays in the GADA titers assigned to a subset of T1D sera (IDS285, IDS301, IDS319, IDS325) (Mann Whitney test: \( P < 0.0001, P < 0.0011, P < 0.0001, P < 0.0008 \), respectively) (Figure 5C).
Discussion

Workshops aimed at the standardization of type 1 diabetes (T1D) associated autoantibodies were originally established by an international community of researchers trying to address early discrepancies in results between laboratories measuring Islet Cell Antibodies (28,29). The success of this initiative led to the recognition of the crucial role of assay format and protocols in determining the results of T1D immunoassays (30) and to the implementation of a harmonization and assessment/validation program for all major T1D autoantibodies.

Following the molecular identification of GAD65 as a major T1D autoantigen, harmonization workshops of GADA measurement accompanied the continuous development, testing and validation of a variety of GADA immunoassays (31). Early interlaboratory comparison studies led to the emergence of the liquid phase immunoprecipitation radio binding assay (RBA) as a widely implemented de facto gold standard for GADA measurement (32).

RBAs not only repeatedly demonstrated good sensitivity and specificity in several interlaboratory comparisons (33) but also allowed laboratories to produce radiolabeled GAD65 using a simple in vitro system and, since in RBA antigen/antibody binding occurs in liquid phase, assuaged the concern for the preservation of GADA conformational epitopes without the disruption often associated with antigen adsorption to a solid phase (34).

The IASP2018 interlaboratory comparison study saw the continued implementation of the RBA format by the majority of laboratories together with an expanded adoption of alternative non-radioactive formats, allowing for a comprehensive comparison between classical RBA and other assay formats in terms of diagnostic, rather than analytical, sensitivity, specificity, accuracy, concordance and performance.
Regarding the analysis of GADA assays performance in IASP2018, in addition to calculating the area under the ROC curve, possibly the most widely used metric of a diagnostic test performance, we evaluated the partial ROC-AUC after imposing a specificity level of 95% (pAUC95). This approach improves one of the major limitations and confounders of using the total ROC-AUC i.e. the inclusion in the analysis of regions of low assay specificity that are not clinically relevant (12).

The characteristics of RBAs submitted to IASP2018, showed the presence of several alternative protocols differing with regard to assay buffers, washing method (i.e. centrifugation vs filtration), amount of test serum and antigen, radiolabel ($^{35}$S, $^{125}$I, $^{3}$H), and the use of full-length or truncated GAD65 antigens (24–26). Unsurprisingly, RBAs presented a large degree of variability in performance, with only a few RBAs using protocols developed in house achieving a pAUC95 above the median of all assays.

Among these protocol variables, the use of truncated GAD65 antigen corresponding to amino acids 96-585 of the full-length protein, was associated with better pAUC95 compared to most RBAs using full-length GAD65 in both RBA and LIPS assays, another immunoprecipitation-based format.

In IASP2018, the second most widely adopted immunoassays were commercial bridge-ELISAs. Bridge-ELISAs showed the most homogeneous and highest pAUC95 of all assays, save for the novel PCR based ADAP format, and proved more sensitive and specific overall than most other assays, with the few positive scores discrepant with those of high-performance RBA and LIPS assays, essentially limited to low titer GADA samples.

We then reassessed positive scores after applying a threshold based on a common pre-defined specificity of 95% to all assays, a threshold that, while not ideal if applied to population screening, facilitated comparisons across this limited sample set. This re-analysis showed that
discrepancies in low titer GADA cases were resolved in the majority of LIPS and RBA assays, suggesting that most liquid phase assays in IASP2018 generally adopted more conservative thresholds compared to the bridge-ELISA. The consistent detection by some RBAs and LIPS assays of slightly increased antibody binding in some specific blood donor sera from the IASP2018 sample set, suggests that threshold selection in these assays might have been driven by the presence of low, non-disease specific antigen binding in some local control samples. Based on prior publications it can be speculated that at least part of this binding might be attributed to low affinity antibodies (25,35).

The comparison of quantitative GADA results in IASP2018, was complicated by the implementation of a variety of local non-standardized arbitrary units and calculation algorithms to express results in place of the international units endorsed by the WHO. Previous workshops addressed this source of inter-assay variability by distributing the WHO standard serum and encouraging reporting of GADA in WHO units. Since the WHO reference consists of a strongly autoantibody positive human serum and constitutes an intrinsically finite resource, its distribution was meant to be used for recalibration of local standards followed by conversion of local to WHO units. However, after encouraging preliminary results (36) further analyses did not confirm the same level of concordance and reproducibility of WHO units across laboratories (37). This prompted the design of a common protocol, including an alternative set of GADA serum standards calibrated against the WHO serum with the aim of harmonizing RBA assays employed by NIDDK sponsored consortia (23).

In IASP2018 the majority of assays reporting results in WHO units consisted of bridge-ELISAs, while only two RBAs using a protocol developed in house did so. NIDDK standards were employed by only three laboratories, and these submitted results for RBAs using the NIDDK harmonized protocol and/or LIPS assays.
Overall, while IASP2018 GADA assays showed a reasonable inter-assay agreement of GADA titer ranks among the assays using WHO units, at least in cases if not in control subjects, we observed a relatively lower concordance of attributed GADA titers, particularly in a subset of cases, despite the use of standardized protocols and centrally prepared calibrators.

Moreover, a dichotomy in GADA titers attributed to selected sera was evident when assays were grouped according to antigen/antibody binding phase into hybrid solid/liquid or just liquid phase assays. While the underlying reason for this behavior remains to be clarified, multiple potential causes can be hypothesized, like epitope alterations following the addition of tags (e.g. biotin residues in the bridge-ELISA or luciferase enzyme in LIPS), subtle differences in primary amino acid sequence, post translational-modifications in different expression systems, and the adsorption of antigens onto solid surfaces in the bridge-ELISA. The frequency of this phenomenon remains to be ascertained along with its potential impact, if any, on autoantibody-based screening strategies, which are currently mostly based on RBAs.

Several novel immunoassays were submitted to IASP2018, a development likely spurred by the continuous legislative and logistic pressure against the use of radioactive substances and the expected future implementation of antibody-based population screening programs for T1D. Among these assays, the PCR based ADAP assay achieved both high sensitivity and perfect specificity, while the rest showed variable performance that in some cases was dramatically inferior to that of more mature formats.

In conclusion, the IASP 2018 results depict the field of GADA measurement as both mature, with numerous assays achieving high performance, but also in relative flux, with the active development and deployment of novel immunoassays dispensing altogether with the need for radio-isotopic tracers. We believe that these results confirm the usefulness of harmonization programs, not only as providers of unbiased diagnostic performance assessment to participants.
but also as an arena in which research laboratories and companies can learn valuable lessons for improving immunoassays for T1D autoantibodies.
Acknowledgements

We wish to thank all the people and the clinicians who contributed blood samples without whom generosity and effort the IASP workshops and collaborative studies could not take place. The authors acknowledge the support received by IASP study through a TEDDY (38) grant administered at the University of South Florida (Tampa, FL) and by the Type 1 Diabetes TrialNet/IASP Ancillary studies arm (39).

References


Assay Format and Contributes to Differences in the Specificity of GAD Autoantibody Assays for Type 1 Diabetes. Diabetes 2015;64:3239–46.


<table>
<thead>
<tr>
<th>Format</th>
<th># assays</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
<th>ROC-AUC(^a)</th>
<th>pAUC95(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antibody Dependent Agglutination PCR (ADAP)</td>
<td>1</td>
<td>80</td>
<td>100</td>
<td>92.9</td>
<td>0.96</td>
<td>0.041</td>
</tr>
<tr>
<td>Bridge-ELISA</td>
<td>12</td>
<td>78.0 (68.0, 80.0)</td>
<td>100.0 (95.5, 100.0)</td>
<td>91.4 (88.6, 92.8)</td>
<td>0.89 (0.88, 0.94)</td>
<td>0.039 (0.036, 0.041)</td>
</tr>
<tr>
<td>Radio Binding Assay (RBA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truncated GAD65 aa96-585</td>
<td>4</td>
<td>69.0 (66.0, 74.0)</td>
<td>100.0 (95.6, 100.0)</td>
<td>88.2 (87.9, 89.3)</td>
<td>0.84 (0.82, 0.88)</td>
<td>0.037 (0.036, 0.039)</td>
</tr>
<tr>
<td>Local protocol</td>
<td>9</td>
<td>66.0 (50.0, 72.0)</td>
<td>97.8 (95.7, 100.0)</td>
<td>87.1 (82.1, 88.6)</td>
<td>0.84 (0.75, 0.89)</td>
<td>0.034 (0.029, 0.037)</td>
</tr>
<tr>
<td>Harmonized protocol</td>
<td>5</td>
<td>64.0 (58.0, 82.0)</td>
<td>98.9 (94.4, 100.0)</td>
<td>85.0 (81.4, 92.9)</td>
<td>0.87 (0.77, 0.91)</td>
<td>0.032 (0.025, 0.039)</td>
</tr>
<tr>
<td>Luciferase Immuno Precipitation System (LIPS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NLuc reporter GAD65 aa96-585</td>
<td>4</td>
<td>70.0 (60.0, 76.0)</td>
<td>99.4 (98.9, 100.0)</td>
<td>88.9 (85.7, 90.7)</td>
<td>0.87 (0.86, 0.89)</td>
<td>0.036 (0.036, 0.037)</td>
</tr>
<tr>
<td>NLuc reporter GAD65 aa188-585</td>
<td>2</td>
<td>71.0 (68.0, 74.0)</td>
<td>98.3 (97.8, 98.9)</td>
<td>88.6 (87.9, 89.3)</td>
<td>0.86 (0.86, 0.86)</td>
<td>0.035 (0.034, 0.035)</td>
</tr>
<tr>
<td>NLuc reporter GAD65 aa1-585</td>
<td>1</td>
<td>72.0</td>
<td>98.9</td>
<td>89.3</td>
<td>0.84</td>
<td>0.034</td>
</tr>
<tr>
<td>FLuc reporter GAD65 aa96-585</td>
<td>1</td>
<td>54</td>
<td>96.7</td>
<td>81.4</td>
<td>0.79</td>
<td>0.019</td>
</tr>
<tr>
<td>Electro Chemi Luminescence (ECL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECL</td>
<td>2</td>
<td>78.0 (78.0, 82.0)</td>
<td>97.2 (95.6, 98.9)</td>
<td>91.1 (89.3, 92.8)</td>
<td>0.89 (0.88, 0.91)</td>
<td>0.036 (0.035, 0.038)</td>
</tr>
<tr>
<td>ECL IgG specific</td>
<td>1*</td>
<td>66.0</td>
<td>96.7</td>
<td>85.7</td>
<td>0.88</td>
<td>0.030</td>
</tr>
<tr>
<td>ECL IgM specific</td>
<td>1*</td>
<td>30.0</td>
<td>97.8</td>
<td>73.6</td>
<td>0.70</td>
<td>0.016</td>
</tr>
<tr>
<td>ECL IgA specific</td>
<td>1*</td>
<td>28.0</td>
<td>97.8</td>
<td>72.9</td>
<td>0.67</td>
<td>0.012</td>
</tr>
<tr>
<td>Chemi Luminescence Immuno Assay (CLIA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiplex Plasmonic Near-InfraRed Fluorescence (MPNIRF)</td>
<td>2</td>
<td>72.0 (68.0, 76.0)</td>
<td>96.7 (93.3, 100.0)</td>
<td>87.9 (84.3, 91.4)</td>
<td>0.88 (0.85, 0.92)</td>
<td>0.035 (0.031, 0.039)</td>
</tr>
<tr>
<td>Luminex Bead Immunooassay (LBI)</td>
<td>2</td>
<td>36.0 (36.0, 36.0)</td>
<td>87.2 (86.7, 87.8)</td>
<td>68.9 (68.6, 69.3)</td>
<td>0.58 (0.57, 0.59)</td>
<td>0.011 (0.011, 0.012)</td>
</tr>
<tr>
<td>Multiplex Fluorescence Energy Transfer (MFRET)</td>
<td>1</td>
<td>2.0</td>
<td>97.8</td>
<td>63.6</td>
<td>0.18</td>
<td>0.001</td>
</tr>
<tr>
<td>All assays - median (IQR)</td>
<td>48</td>
<td>69.0 (64.0, 76.0)</td>
<td>98.9 (96.7, 100.0)</td>
<td>88.6 (84.5, 90.7)</td>
<td>0.87 (0.83, 0.89)</td>
<td>0.036 (0.032, 0.039)</td>
</tr>
</tbody>
</table>

\(^a\) Area under the Receiver Operator Characteristics curve; \(^b\) partial ROC-AUC imposing a specificity greater than 95%; *Multiplexed ECL assay
Figure legends

Figure 1:
Scatter plots of sensitivity and specificity of GADA assays based on laboratory assigned GADA positive or negative scores for 50 cases and 90 controls. Filled circles stand for individual assays. Dashed lines mark the median sensitivity and specificity of all assays. Assays are categorized according to format and its variants. Categories are sorted by their median assay performance.

Figure 2:
Performance of GADA assays in IASP 2018. Shown are the partial areas under the ROC curve at 95% specificity (pAUC95) of each assay (grey filled dots) and the probability density estimates of the pAUC95 distribution. Assays are grouped by format and its variants and the groups sorted according to their median pAUC95. The dashed line marks the median pAUC95 of all assays.

Figure 3:
Tilemap of GADA positive (dark grey) or negative (light grey) scores assigned by laboratories to cases. Assays are grouped by format and its variants and the groups sorted according to their median pAUC95. Cases are ordered by median GADA titer rank. A cumulative score of individual Ig class results was added for lab 1306 in the ECL group. Cases with distinctive patterns discussed in the text are indicated by a black triangle.

Figure 4:
Tilemap of GADA positive (dark grey) or negative (light grey) scores assigned by laboratories to controls. Assays are grouped by format and its variants and the groups sorted according to
their median pAUC95. Omitted are 40 controls negative in all assays. A cumulative score of 40 individual Ig class results was added for lab 1306 in the ECL group. Controls with distinctive patterns discussed in the text are indicated by a black triangle.

Figure 5:

Analysis of GADA titers concordance in IASP2018. (A), GADA titer ranks of each case in different assays (circles, filling is proportional to pAUC95). The black line shows the median rank of all assays. (B), Boxplots of GADA WHO units in each case. Assays are grouped by binding phase (dark grey: hybrid solid/liquid all bridge-ELISA assays; light grey: liquid, include ECL, LIPS and RBA assays). (C), Enlargement showing samples with significant differences in WHO units between assay groups. Cases selected in panel C are indicated in panels A and B by a black triangle.
ADAP: antibody dependent agglutination PCR
CLIA: Chemi Luminescence Immuno Assay
ECL: Electro Chemi Luminescence assay
ECL-IgG: IgG specific ECL assay
ECL-IgA: IgA specific ECL assay
ECL-IgM: IgM specific ECL assay
LBI: Luminex Bead Immunoassay
LIPS GAD96-585: Luciferase Immuno Precipitation System using truncated GAD65 aa 96-585
LIPS GAD188-585: Luciferase Immuno Precipitation System using truncated GAD65 aa 188-585
MFRET: Multiplex Fluorescence Resonance Energy Transfer
MPNIRF: Multi Plasmonic Near Infra Red Fluorescence
RBA GAD96-585: Radio Binding Assay using truncated GAD65 aa 96-585
RBA GAD188-585: Radio Binding Assay using truncated GAD65 aa 188-585
RBA local: RBA using protocol developed in house
RBA NIDDK: RBA using the NIDDK harmonized protocol
Figure 2:

Performance of GADA assays in IASP 2018. Shown are the partial area under the ROC curve at 95% specificity (pAUC95) of each assay (grey filled dots) and the probability density estimates of the pAUC95 distribution. Results are grouped according to assay format and variants and sorted by group specific median pAUC95. The dashed line marks the overall median pAUC95.