Antinuclear antibody prevalence in a general pediatric cohort from Mexico City: discordance between immunofluorescence and multiplex assays

Emily C Somers¹ ³
Seetha U Monrad¹
Jeffrey S Warren⁴
Maritsa Solano⁵
Lourdes Schnaas⁶
Mauricio Hernandez-Avila⁵
Martha Maria Tellez-Rojo⁵
Howard Hu⁷

¹Division of Rheumatology, Department of Internal Medicine, ²Department of Environmental Health Sciences, ³Division of Obstetrics & Gynecology, ⁴Division of Clinical Pathology, Department of Pathology, University of Michigan, Ann Arbor, MI, USA; ⁵Center for Nutrition and Health Research, National Institute of Public Health, Cuernavaca, Morelos, ⁶Department of Developmental Neurobiology, National Institute of Perinatology, Mexico City, Mexico; ⁷Occupational and Environmental Health, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

Objective: To characterize antinuclear antibody (ANA) prevalence according to distinct assay methodologies in a pediatric cohort from Mexico City, and to further examine associations with age and sex.

Methods: Serum ANA were measured by indirect immunofluorescence assay (IFA) and multiplex immunoassay in 114 children aged 9–17 years. IFA was considered positive at a cutoff titer of ≥1:80. Agreement between assay methods was assessed by kappa statistic. Sensitivity, specificity, and 95% confidence intervals (CIs) of the multiplex were computed with IFA as the reference standard.

Results: Of the 114 children (mean age 14.7 [standard deviation 2.1] years; 54 [47%] female), 18 of 114 (15.8%) were ANA positive by IFA, and 11 of 114 (9.6%) by 11-antigen multiplex assay. ANA prevalence was higher in females compared with males by both of the methods (ratios 1.6–1.9 to 1). Agreement between tests was classified as slight by kappa (κ=0.177 [95% CI −0.051, 0.406]). The multiplex immunoassay had sensitivity of 22.2% (95% CI 6.4, 47.6) and specificity of 92.7% (95% CI 85.6, 97.0), and failed to capture 3 of 4 (75%) of the high-titer (≥1:1280) IFA-positives.

Conclusion: Up to 15% of children in this general population cohort were ANA positive, with a higher rate of positivity among females according to both assay methods. Substantial discordance in ANA results was found between IFA and multiplex methods, even for high-titer IFA positives. These findings underscore the need to sufficiently account for assay characteristics when interpreting ANA test results, and support IFA as the more appropriate assay for studies of subclinical autoimmunity.

Keywords: autoreactivity, biomarker, immune dysfunction, preclinical, subclinical autoimmunity, autoantibodies, pediatric, epidemiology

Introduction
Antinuclear antibodies (ANAs), increasingly referred to by the more technically correct term “autoantibodies to cellular antigens”,¹ are associated with numerous autoimmune diseases, such as systemic lupus erythematosus (SLE), systemic sclerosis, Sjögren’s syndrome, and idiopathic inflammatory myopathies. ANAs are highly sensitive for some of these disorders, and are considered the serologic hallmark of SLE, with 95% of SLE patients being ANA-positive.²

Numerous studies have demonstrated that circulating autoantibodies can be detected several years in advance of clinical autoimmune diagnoses. A seminal study of 130 SLE patients found that 88% had at least one SLE-associated autoantibody up...
to several years prior to diagnosis. The existence of auto-
antibodies in asymptomatic individuals who subsequently
develop disease has been reported in a variety of other auto-
immune conditions, including rheumatoid arthritis and
type 1 diabetes. The terms “preclinical”, “subclinical”, and
even “benign” autoimmunity have been used to describe the
presence of autoantibodies or other markers of autoreactivity
or altered immune response in the absence of tissue injury
or clinical manifestations. We favor the term “subclinical
autoimmunity” in prospective settings, as a reflection that
not all individuals with markers of autoimmunity will prog-
ress to overt autoimmune disease. A conceptual framework
has emerged, whereby a subclinical autoimmunity phase of
variable duration is thought to represent an important early
stage in the development of clinical autoimmune disease.
Accurate identification of which individuals will progress
from subclinical autoimmunity to a pathologic state, and
identification of underlying risk factors, depends on the abil-
ity to reliably measure autoantibodies in the epidemiologic
setting.

Approximately 13% of the general population demon-
strates ANA positivity according to standard immunofluores-
cence methodology (at a 1:80 cutoff titer). While pediatric
populations have been less well studied, a US-NHANES
study reported ANA prevalence rates among 12–19-year-olds
of approximately 9.5% for non-Hispanic whites, 12.5% for
study reported ANA prevalence rates among 12–19-year-olds
populations have been less well studied, a US-NHANES
variable duration is thought to represent an important early
stage in the development of clinical autoimmune disease. A
study.16

Methods

Study population

The study population was comprised of offspring born
to mothers who had previously enrolled in a birth cohort
in Mexico City – the Early Life Exposures in Mexico
to Environmental Toxicants (ELEMENT) study – dur-
ing 1994–2003. As described in detail elsewhere, the
original ELEMENT study enrolled low-to-moderate
income, pregnant women from the National Institute of
Perinatology, Hospital General Dr. Manuel Gea Gonzale-
zez, or clinics affiliated with the Mexican Social Security
Institute. Eight hundred twenty-eight children of these
mothers were enrolled during 2007–2011 into a follow-up
study. For the present study, 114 offspring with banked
serum from follow-up visits in 2010–2011 were included.
Ethics approval was obtained from the respective institu-
tional review boards (National Institute of Public Health
of Mexico, Harvard School of Public Health, University
of Michigan, and participating hospitals). Maternal writ-
ten informed consent and written assent from the children
were obtained.

ANA testing

ANA assays were performed in the CLIA (Clinical Labora-
tory Improvement Amendments)-certified Clinical Immun-
ology Laboratory at the University of Michigan Health
System using stored serum from the ELEMENT biobank.
Specimens were initially collected in standard gold SST
(serum separator tube) vacutainers and allowed to clot at
room temperature before centrifugation at 3,500 rpm for 10
minutes. Serum was then aliquoted into cryovials and stored
at −70°C prior to analysis.

Indirect IFA

Serum was applied to HEp-2 cell slides (Fluorescent HEp-2
ANA; Immuno Concepts, Sacramento, CA, USA) after
dilution in 0.15 M NaCl and 10 mM phosphate buffered
saline (PBS), pH 7.4, and incubated for 30 minutes at room
temperature in a moist chamber. After washing twice
in PBS for 10 minutes, cells were incubated with fluorescein
isothiocyanate-conjugated goat anti-human Ig (IgG heavy
and light chains) for 30 minutes in the dark. After two further
washes, slides were assembled with buffered glycerol, pH
9.5, and coverslipped. ANA titer was determined by testing
successive twofold dilutions of the serum to endpoint. PBS
buffer and negative human control serum supplied with the
commercial kit (Immuno Concepts) were used as negative controls. Analysis was performed by two independent expert observers using a Zeiss Rx 100 microscope under 400× magnification. Samples were classified as ANA-HEp-2 positive if a well-defined indirect immunofluorescence pattern was identified (≥3+/4) at 1:80 dilution by both observers. Patterns were determined based on standard classification criteria. Fluorescence photomicrographs from our laboratory with representative images are included in Figure 1A–D. Expert consensus is lacking as to the titer to consider positive for children, and some have used a dilution of 1:40.1 We utilized a cutoff of 1:80, as it is our standard institutional positive/negative cutoff based on a laboratory validation study in which the false-positive rate was set at ≤5% (J Warren, unpublished data, 2007).

Multiplex immunoassay
We used the ANA multiplex immunoassay (BioPlex 2200 ANA Screen; BioRad, Hercules, CA, USA), an automated two-color fluorescence microbead flow system in which recombinant or purified native antigen has been linked to beads. Specific antigen beads were incubated with serum, followed by incubation with fluorochrome-linked antibody directed against human immunoglobulin G. This test detects autoantibodies to 11 antigens with clinical relevance to autoimmune/rheumatic diseases (double-stranded DNA [dsDNA], SS-A, SS-B, Sm, RNP, SmRNP, Scl-70, Jo-1, centromere B, ribosomal P, chromatin). A negative result indicates that the serum shows no reactivity for these antibodies.

Statistics
Characteristics of the study population were compared between groups using two-sample t-tests or Wilcoxon rank-sum tests for continuous data, or chi-squared tests for categorical data. Agreement between assays was assessed by the kappa statistic. Kappa statistic interpretation was classified as follows: <0, less than chance; 0.01–0.20, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; 0.81–0.99, almost perfect.26 Sensitivity and specificity, with exact binomial 95% confidence intervals (CIs), of the multiplex test were calculated, with IFA as the “gold standard”. Stata v.13 software was used (StataCorp LP, College Station, TX, USA).

Results
Of the 114 children in this study, 54 (47%) were female. Ages at follow-up ranged from 9 to 17 years (mean [SD] 14.7 [2.1]), and were similar between males and females (female 14.5 [0.30]; male 15.0 [0.26]; p=not significant).

ANA prevalence
ANA results according to assay methodology are presented in Table 1. ANA positivity was found in 18 of 114 (15.8%) by IFA and 11 of 114 (9.6%) by multiplex. Three of the children who were multiplex positive had only dsDNA positivity among the 11 antigens tested. The presence of only anti-dsDNA by multiplex, in the absence of other relevant findings, is considered by some experts to represent a false-positive, often due to anti-single-stranded DNA (ssDNA) antibodies.27 When these three children were classified as negative by multiplex, the multiplex-positive proportion decreased to 8 of 114 (7%) overall, 5 of 54 for females (9.3%), and 3 of 60 for males (5%). Regardless of assay methodology, in this pediatric population, a higher proportion of females compared with males tested positive for ANAs, though not reaching statistical significance, with female-to-male ratios for ANA positivity ranging from 1.6–1.9 to 1. Speckled immunofluorescence patterns predominated at lower titers, and the homogeneous pattern at higher titers (Figure 2A). One mitotic spindle pattern was found, which is considered an atypical pattern and thus classified as ANA-negative at our institution; this person also was negative by multiplex. No other atypical patterns were observed in our cohort, including

Figure 1
Representative photomicrographs of antinuclear antibody (ANA) patterns at 400x magnification. (A) negative ANA; (B) positive/homogeneous pattern; (C) positive/speckled pattern; (D) positive/nucleolar pattern.
antibodies to dense fine speckles (anti-DFS70) which have been increasingly noted in healthy populations. From the multiplex-11 panel, 4 specificities did not yield any positives (ribosomal P, SS-A, centromere B, and Jo-1); distribution shown in Figure 2B.

**Age patterns**

According to IFA, mean age was similar between those who screened positive (14.7 years [SD 2.2]) versus negative (15.2 years [1.9]); $p$=not significant. By multiplex, there was a trend toward older age among multiplex-positives (15.8 years [1.3]) compared with multiplex-negatives (14.6 years [2.2]); $p$=0.07 by Wilcoxon rank-sum test. In general, the IFA screen captured positives from a broader age distribution, with the multiplex predominantly capturing positives among those aged >15 years (Figure 3).

**Agreement between assays**

The observed agreement between IFA and multiplex-11 in the overall study population was 81.6%, compared to expected agreement by chance of 77.6%, yielding a kappa coefficient of $\kappa=0.177$ (95% CI $-0.051$, 0.406); by conventional interpretation, this indicates “slight” agreement. Agreement was higher for boys ($\kappa=0.305$ [95% CI $-0.075$, 0.684]) than for girls ($\kappa=0.076$ [95% CI $-0.208$, 0.360]). The extent of overlap between tests is presented Figure 4. Of the 25 children with a

Table 1 ANA positivity according to assay method in a pediatric cohort (aged 9–17 years) from Mexico City

<table>
<thead>
<tr>
<th></th>
<th>Overall, n=114, No (%)</th>
<th>Female, n=54, No (%)</th>
<th>Male, n=60, No (%)</th>
<th>$P$-value (female vs male)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ANA positivity, by assay method</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IFA ($\geq$1:80)</td>
<td>18 (15.8)</td>
<td>11 (20.4)</td>
<td>7 (11.7)</td>
<td>NS</td>
</tr>
<tr>
<td>Multiplex-11</td>
<td>11 (9.6)</td>
<td>7 (13.0)</td>
<td>4 (6.7)</td>
<td>NS</td>
</tr>
<tr>
<td>Multiplex-10 (excludes dsDNA+)</td>
<td>8 (7)</td>
<td>5 (9.3)</td>
<td>3 (5.0)</td>
<td>NS</td>
</tr>
<tr>
<td><strong>Titers among IFA+</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:80</td>
<td>6 (33.3)</td>
<td>5 (45.5)</td>
<td>1 (14.3)</td>
<td>–</td>
</tr>
<tr>
<td>1:160</td>
<td>4 (22.2)</td>
<td>1 (9.1)</td>
<td>3 (42.9)</td>
<td>–</td>
</tr>
<tr>
<td>1:320</td>
<td>4 (22.2)</td>
<td>2 (18.2)</td>
<td>2 (28.6)</td>
<td>–</td>
</tr>
<tr>
<td>1:1280</td>
<td>2 (11.1)</td>
<td>2 (18.2)</td>
<td>0 (0)</td>
<td>–</td>
</tr>
<tr>
<td>1:2560</td>
<td>2 (11.1)</td>
<td>1 (9.1)</td>
<td>1 (14.3)</td>
<td>–</td>
</tr>
<tr>
<td><strong>Fluorescence patterns</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homogeneous</td>
<td>8 (44.4)</td>
<td>4 (36.4)</td>
<td>4 (57.1)</td>
<td>NS</td>
</tr>
<tr>
<td>Nucleolar</td>
<td>1 (5.6)</td>
<td>0 (0)</td>
<td>1 (14.3)</td>
<td>–</td>
</tr>
<tr>
<td>Speckled</td>
<td>9 (50)</td>
<td>7 (63.6)</td>
<td>2 (28.6)</td>
<td>–</td>
</tr>
<tr>
<td><strong>Number of Ag+ specificities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>8 (7)</td>
<td>5 (9.3)</td>
<td>3 (5)</td>
<td>NS</td>
</tr>
<tr>
<td>2</td>
<td>2 (1.8)</td>
<td>2 (3.7)</td>
<td>0 (0)</td>
<td>–</td>
</tr>
<tr>
<td>5</td>
<td>1 (0.9)</td>
<td>0 (0)</td>
<td>1 (1.7)</td>
<td>–</td>
</tr>
</tbody>
</table>

Notes: *Among IFA positives, *among multiplex-11 positives.

Abbreviations: Ag, antigen; ANA, antinuclear antibody; dsDNA, double-stranded DNA; IFA, immunofluorescence assay; NS, not significant.

Figure 2 (A) ANA titers and patterns, by indirect IFA. (B) Autoantibody specificities by multiplex-11 assay. Four specificities did not yield any positives (ribosomal P, SS-A, centromere B, and Jo-1); distribution shown in Figure 2B.

Abbreviations: ANA, antinuclear antibody; dsDNA, double-stranded DNA; IFA, immunofluorescence assay.
positive ANA by either test, only 4 (16%) were dual-positive (by both IFA and multiplex); the lower overall prevalence captured by multiplex did not simply reflect a subset of the IFA positives. All four dual-positives occurred in children older than 15 years. With IFA (≥1:80) as the reference standard, multiplex sensitivity was 22.2% for both the 11- and 10-antigen panels; specificity was 92.7% and 95.8%, respectively. All three of the multiplex dsDNA-only positives were negative by IFA, lending support for exclusion of the dsDNA antigen from this test. There was discordance across the range of IFA titers (Table 2), and the multiplex failed to capture 3 of 4 (75%) of the high-titer IFAs (≥1:1280).

**Table 2** Comparison of IFA titers and number of multiplex-positive antigens

<table>
<thead>
<tr>
<th>Multiplex</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>dsDNA+ only</td>
</tr>
<tr>
<td>IFA titer</td>
<td></td>
</tr>
<tr>
<td>&lt;1:80</td>
<td>89</td>
</tr>
<tr>
<td>1:80</td>
<td>4</td>
</tr>
<tr>
<td>1:160</td>
<td>4</td>
</tr>
<tr>
<td>1:320</td>
<td>3</td>
</tr>
<tr>
<td>1:1280</td>
<td>2</td>
</tr>
<tr>
<td>1:2560</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>103</td>
</tr>
</tbody>
</table>

**Discussion**

We have described ANA prevalence according to competing assay methodologies in a pediatric cohort from the general Mexico City population. Understanding serologic patterns associated with subclinical autoimmunity in various population subsets is an important component of immunoepidemiology. As ANA results are highly dependent upon assay type, appreciation of methodologic differences and their implications is essential. Further, as significant health disparities in the incidence, severity and age of onset have been documented for autoimmune diseases such as SLE,29–32 it follows that subclinical autoimmunity should be specifically characterized in various population demographics. Among the 9–17-year-olds in our study, 15.8% were positive by IFA and 9.6% by multiplex. This rate of ANA-IFA positivity extends the high end of the range of approximately 9.5%–14.6% reported in other pediatric populations using IFA at a 1:80 cutoff,14–16 and is compatible with epidemiologic data indicating that risk of SLE for Hispanics is higher than for non-Hispanic whites, yet lower than for blacks.33,34 In our study, the proportion of positive females was close to double that of males for each assay method. It is well-recognized that autoimmune diseases disproportionately affect females, with female-to-male prevalence ratios of up to 9:1.35 Our data add to accumulating evidence that the phenomenon of subclinical autoimmunity is likewise more prevalent among females, even in childhood and adolescence.36

We found that ANA positivity was lower based on the multiplex methodology compared to IFA, which is consistent with findings from patients with clinical autoimmune diseases such as SLE and other ethnic groups.19 This is in part due to the fact that solid-phase methods such as multiplex assays test for a restricted range of approximately 8–11 antigens, compared with the broader range of 100–150 possible autoantigens identified by IFA.37 However, we found
that the lower positivity rate for the multiplex assay was not simply a reflection of a smaller subset of the IFA positives being captured by multiplex, as would have been theoretically expected. Rather, discordance occurred bidirectionally (ie, multiplex positive/IFA negative and vice versa); only 4 specimens among the 114 tested were concordant positives (yielding an ANA positivity rate of 3.5% for specimens classified as positive by both methods). Such discordance may be due to recombinant antigens, such as those used in the multiplex assay, displaying different epitopes than those on native antigens, such as are used in the IFA. Further, autoantigen profiles have been shown to vary between racial and ethnic groups; thus, the defined set of antigens on the multiplex may not be uniformly applicable across groups. In our study, the discordance between multiplex and IFA was particularly notable at high titers of IFA positivity, with three quarters of IFA positives at titers of 1:1280 or higher being identified as negative by multiplex.

Moreover, our data demonstrate that the performance characteristics of the 2 ANA assays varied between boys and girls, with lower agreement among girls, where the coefficient of agreement was classified as “slight”. This may indicate that there is a broader diversity of autoantigens among girls that is captured by the IFA (which detects all autoantibodies binding to Hep-2 cell nuclei) but not by multiplex (with its restricted range of available epitopes). The lower agreement among girls is concerning given the substantially higher risk of autoimmune disease among females. We also found better agreement among older children.

While we could not directly examine differences across race or ethnicity in this Mexican population, other studies have found the relative performance of ANA assay methods to vary according to race. Bruner et al reported that in both SLE cases and controls from a multiethnic population including African-, Mexican-, and European-Americans, ANA prevalence according to IFA vs multiplex was most similar for African Americans, and least similar for European Americans; among controls, both tests yielded a prevalence rate of ~20% among African Americans, whereas prevalence was 1.3- and 1.9-fold higher for IFA compared to multiplex among Mexican Americans and European Americans, respectively. Further, they found the false-positive rate for multiplex (ie, multiplex positive but IFA negative) among SLE cases to range from 2.1% for African Americans to 7.6% for Mexican Americans, whereas among controls, false positive rates were more similar between groups (11.4%-13.2%). The differential levels of agreement between tests across demographic groups need to be taken into consideration when evaluating choice of screening tool and interpreting results.

Based on the high discordance between tests, and particularly that the multiplex failed to capture the majority of high titer IFA positives in our study population, we would argue that the IFA remains the test of choice for screening, including in the context of epidemiologic research related to subclinical autoimmunity. Despite existing guidelines maintaining support for IFA as the reference standard, the multiplex assay has become increasingly popular due to its automated nature and lack of reliance on specialized technicians, potentially lower cost, and focus on a subset of antigens clinically relevant to autoimmune rheumatic diseases. These attributes are likewise attractive for epidemiologic research. However, we contend that given the low sensitivity (22%) of the multiplex, it is poorly suited as a first-stage or screening test, as it would miss a high proportion of “true positive” cases for whom a positive ANA would prompt further evaluation in the clinical setting, or for whom a signal of immune dysfunction would be missed in the research setting. Conversely, “false positives” would result in the wrong subset of persons being identified as having evidence of autoimmunity. Thus, the multiplex may not ultimately be cost-effective if it triggers unnecessary evaluation or if it misses opportunities for early intervention. Perhaps the multiplex is better suited as a follow-up to a positive IFA screen when identification of antigen specificity is a primary concern.

Understanding patterns of subclinical autoimmunity in the population is important for the elucidation of factors involved in initial perturbation of the immune system and subsequent progression toward a clinical state. While conventional wisdom has been that ANA positivity in children may often be due to infection, Allen et al did not find historical or intercurrent viral infections to be associated with an increased risk of ANA positivity. Studies have begun to explore the association between environmental toxicants and subclinical autoimmunity. We recently reported that in the general US population, low-level methylmercury exposure was associated with higher risk of ANA positivity (by IFA) among females of reproductive age. Further research into the role of environmental exposures in the development of autoimmunity relies on assays with favorable reliability and with broad applicability to various population subsets.

Our study clearly demonstrates that ANA testing based on multiplex and IFA assays yields not only divergent overall prevalence rates but differential agreement across demographic subsets. By simultaneously testing the same specimens by both methods, discordance in results cannot
be ascribed to temporal physiologic changes, as might be contemplated in the clinical setting when serial testing yields opposing results. While the majority of data related to ANA testing stem from clinical series of persons with known or suspected autoimmune disease, we have extended the literature by focusing on a general population cohort of children. Given the homogeneous nature of our study population, the generalizability of results to adults or other racial and ethnic groups is unknown. However, we have focused on an ethnic group with heightened risk of SLE and potentially other autoimmune diseases compared to the non-Hispanic white population, and provide enhanced characterization of ANA testing in this group.

Conclusion
We have demonstrated substantial variation in the performance of competing assays for ANA detection in a pediatric cohort. Our data support a growing body of evidence demonstrating that, despite these two tests often being used interchangeably, they are in fact disparate assays. As in the clinical setting, this should be taken into consideration in research studies of subclinical autoimmunity.

Acknowledgments
This work was supported by NIH/NIEHS grants K01ES019909, P30ES017885, R01ES021446, R01ES007821, P01ES022844 and the National Institute of Public Health/Ministry of Health of Mexico. The American British Cowdray Hospital provided facilities used for subject recruitment and sample collection. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

Disclosure
The authors report no conflicts of interest in this work.

References


