Serology assays used in SARS-CoV-2 seroprevalence surveys worldwide: a systematic review and meta-analysis of assay features, testing algorithms, and performance

CITF Scientific Meeting: 09 March, 2023

Presenters: Rahul K. Arora, Harriet Ware

Ma X, Li Z, …, ***Arora RK****. Vaccines*. 2022. https://doi.org/10.3390/vaccines10122000

SeroTracker

SeroTracker

DISCLAIMER

Funding:

- Public Health Agency of Canada / CITF (this work)
- World Health Organization (this work)
- Robert Koch Institute, Canadian Medical Association Joule, Open Philanthropy (other work)

RKA's disclosures (each unrelated to the present work):

- Past employment, Flagship Pioneering
- Past employment, Health Canada
- Past consulting, Bill and Melinda Gates Foundation Strategic Investment Fund
- Minority shareholder, Alethea Medical

SeroTracker

MOTIVATION

- The breadth of serological assays since the beginning of the pandemic is diverse
  - commercial assays vs self-develoeped assays
  - quantitative vs qualitative
  - RDT vs non-RDT

- Assay performance has direct consequences on the validity of a study,
  - sensitivity (Sn.) and specificity (Sp.) vs group's true antibody positivity (seroprevalence)

- Varied intra-assay performance data
  - Sn. and Sp. are not fixed properties — rather, they depend on evaluation method and reference panel
  - Lack of standardization between the methodology for evaluations
  - Biased estimates with statistical adjustment for Sn. and Sp.

[1] Lau EHY, Tsang OTY, Hui DSC, Kwan MYW, Chan WH, Chiu SS, Ko RLW, Chan KH, Cheng SMS, Perera RAPM, Cowling BJ, Poon LLM, Peiris M. Neutralizing antibody titres in SARS-CoV-2 infections. Nat Commun. 2021 Jan 4;12(1):63. doi: 10.1038/s41467-020-20247-4. PMID: 33397909; PMCID: PMC7782739.

SeroTracker

DEFINE EVALUATION SOURCES

Define evaluation sources

Manufacturer:
Biotech group which invented, evaluated and distributed the serological assay

Third-party head-to-head evaluations:
Reference labs conducted large-scale head-to-head evaluations under controlled and reproducible conditions

Independent field evaluations:
A small size of serological sample pretested using an assay before the formal serosurvey rolls out

- Third-party head-to-head evaluations:
  - head-to-head evaluations
  - reference panels mirrors the complexity of antibody detection in real settings
- Independent field evaluations:
  - representative endemic samples of the demographics and endemic pathogens

SeroTracker

OBJECTIVES

- Describe features and usage of serological assays
- Comprehensively compare the performance of these assays across manufacturers, third-party reference labs, and independent investigator evaluations
- Quantitatively assess the influence of assay performance on seroprevalence estimates

SeroTracker

METHODS - DATA COLLECTION

- Design: systematic review and meta-analysis
- Search time: between 1 January 2020 to 19 November 2021
- Source of articles:
  - **Primary**: Medline, EMBASE, Web of Science, and preprints on Europe PMC
  - **Secondary**: Google News, articles submitted to SeroTracker.com, or recommended by experts

- Data extracted
  - **Basic**:
    - product name, manufacturer, country,
    - antibody isotypes detected (IgG, IgM, IgA, total Ab),
    - test type (ELISA, LFIA, IFA, CLIA, neutralization assay, etc.),
    - antibody target (Spike, Nucleocapsid, others), multiplex detection
    - time to result (RDT/non-RDT)
  - **Performance specific**:
    - Sn. and Sp. as reported by manufacturers or developers
    - Sn. and Sp. validation from either (1) third-party lab validation or (2) independent group field validation

SeroTracker

METHODS - ANALYSIS

- **Descriptive**:
  - Basic characteristics of identified assays - at assay level and at study level
  - **Inter-class performance variation:** Median Sn. and Sp. values for the top 50 assays from three evaluation sources are plotted separately against the WHO criteria for emergency use (Sn. >= 90%, Sp. >= 97%)
  - **Intra-class performance variation:** Bland–Altman plots wo compare manufacturer-reported Sn./Sp. with a third party's lab and independently evaluated Sn./Sp. in pairs

- Modeling
  - **Intra-class performance variation (by evaluation source):** mixed-effect beta regression models for Sn. and Sp. with random effects specified for individual serological assays
  - **Bias introduced by mis-reported assay performance.** 1000 simulated scenarios in which observed seroprevalence ranged from 0.0–99.9%. Adjusted prevalence on assay Sn. and Sp. to compare with the "true" prevalence

SeroTracker

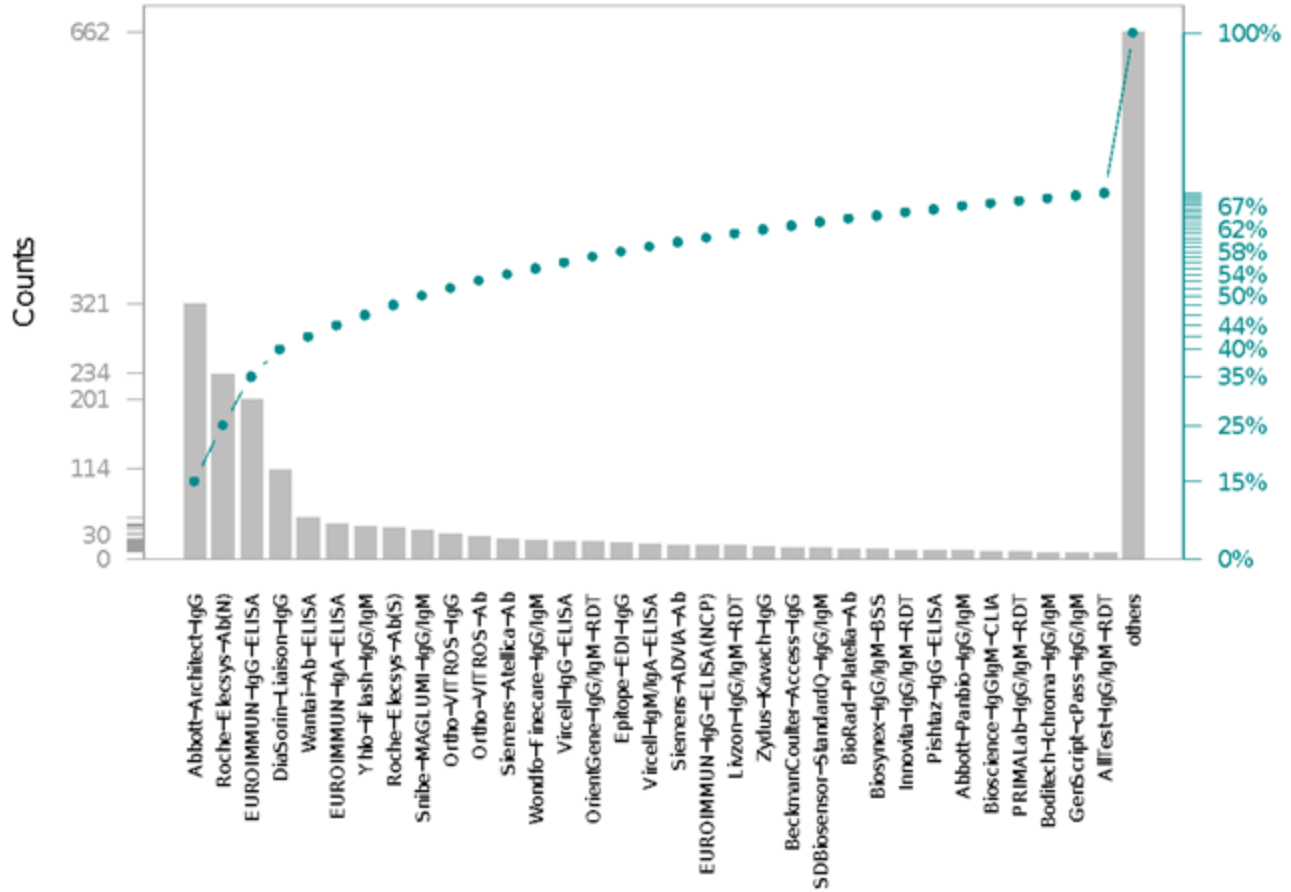
1807 STUDIES USED 572 ASSAYS; TOP 50 ASSAYS USED IN ONLY 67% OF STUDIES

- 1807 source articles with assay data to synthesize:
  - 80.7% of studies used a **single** serological assay (73.1% commercial assays, 18.2% self-developed assays, 8.7% unable to specify)

- 19.3% used a testing algorithm involving **multiple** assays

- 192 commercial serological assays and 380 self-developed assays

- Pareto diagram on assay use frequency: very long tail



SeroTracker

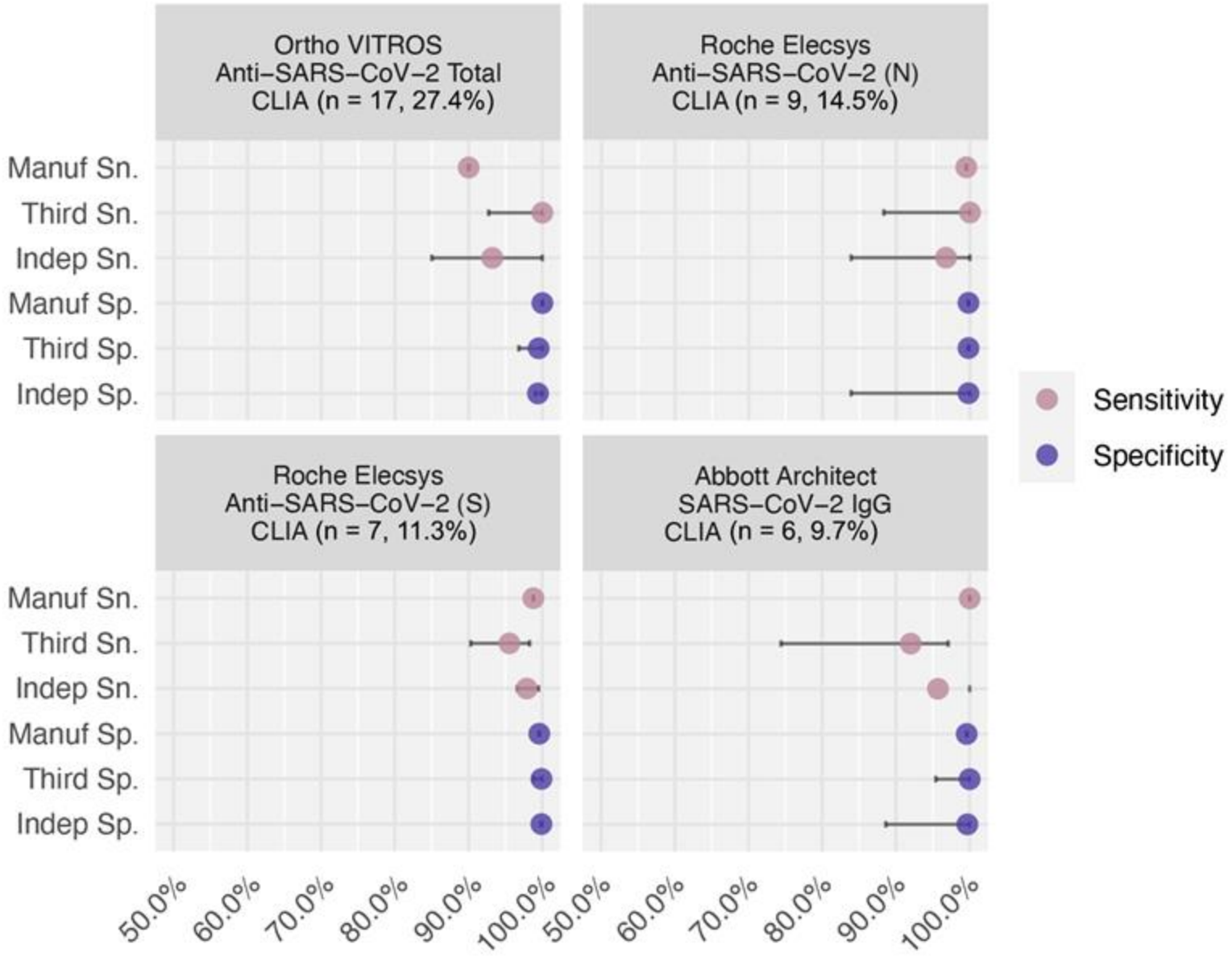
MANUFACTURER REPORTED TEST PERFORMANCE MET EMERGENCY USE CRITERIA FOR 29.7% OF ASSAYS

Assay characteristics	Commercially assays (N = 192)		Self-developed assays (N = 380)	
	n	%	n	%
Developed by				
Manufacturer	162		-	
Lab groups	-		275	
Type of Assays				
ELISA	60	31.3	261	68.7
LFIA	75	39.1	0	0.0
IFA	5	2.6	17	4.5
CLIA (Including CGIA, CMIA)	30	15.6	3	0.8
Neutralization Assay	0	0.0	52	13.7
Others/ Not specified	22	11.5	47	12.4
WHO regions of development				
Africa	0	0.0	12	3.2
America	49	25.5	152	40.0
Eastern Mediterranean	5	2.6	15	3.9
Europe	75	39.1	163	42.9
South-East Asia	4	2.1	6	1.6
Western Pacific	58	30.2	32	8.4
Not Reported	1	0.5	0	0.0
Feature of Assays				
RDT	103	53.6	17	4.5
Non-RDT	89	46.4	363	95.5
Antibody Targets				
Spike	55	28.6	48	12.6
Nucleocapsid	37	19.3	37	9.7
Multiplex Targets (*)	38	19.8	171	45.0
Unknown	62	32.3	124	32.6
Isotypes				
IgG-only	52	27.1	149	39.2
IgG and IgM	103	53.6	31	8.2
Total Antibody (IgG, IgM, IgA)	22	11.5	38	10.0
Other Combinations (**)/ Not Reported	15	7.8	162	42.6
Assay Sn. and Sp.				
Manufacturer/developer reported	91	47.4	124	32.6
Third party validated	118	61.5	-	-
Australia NRL	16	8.3	-	-
Australia Doherty	18	9.4	-	-
US FDA	57	29.7	-	-
FIND Diagnostic	30	15.6	-	-
Netherland CIDC	26	13.5	-	-
Other groups	94	49.0	-	-
Emergency Use (***)				
Yes	57	29.7	-	-
No	135	70.3	-	-

SeroTracker

INTRA-CLASS PERFORMANCE VARIATION - TOP 4 COMMERCIAL ASSAYS IN 62 CANADIAN SEROSURVEYS

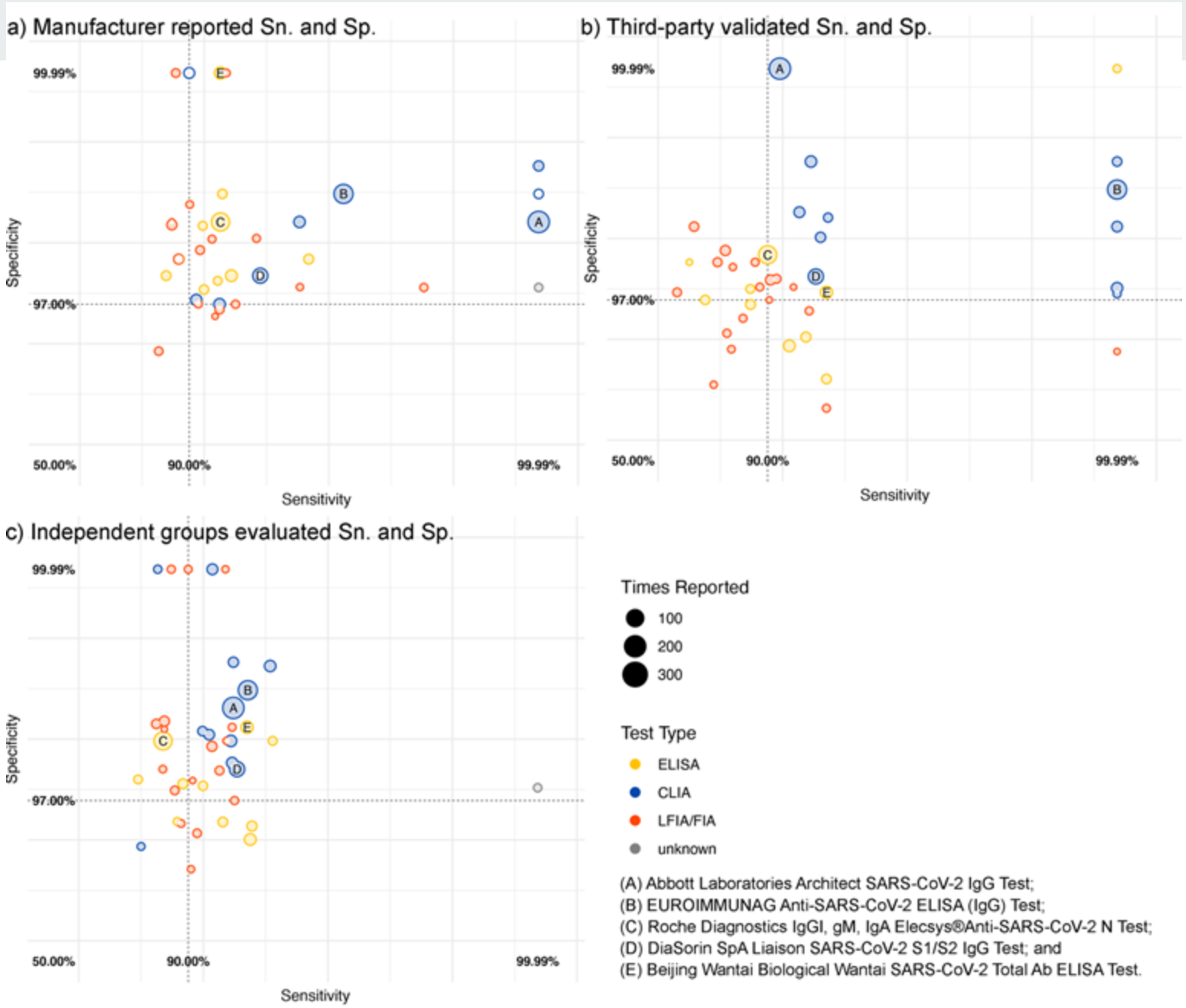
All collected performance data (from Canadian studies and non-Canadian studies) was used to construct intervals



SeroTracker

INTRA-CLASS PERFORMANCE VARIATION - AGAINST WHO EMERGENCY USE CRITERIA FOR TOP 50 COMMERCIAL ASSAYS

Proportion of assays meeting the criteria fell from 76.9% to 46.1% and 53.7% based on third-party and independent evaluations



SeroTracker

Sn. and Sp. were considerably lower according to third parties and independent evaluations

Manufacturer:
averaged Sn. 97.8 (95% CI: 93.9-100)%
averaged Sp. was 99.7 (95% CI: 97.8-100)%

	Sensitivity				Specificity			
›Fixed effects	**Difference in performance against manufacturer value [a] [95% CI]**	**Absolute performance value [b] [95% CI]**	**P [c]**		**Difference in performance against manufacturer value [b] [95% CI]**	**Absolute performance value [a] [95% CI]**	***P* [c]**	
Source of Evaluation								
Manufacturer	*ref.*	93.6% [90.6, 95.7%]	<0.001	*	*ref.*	98.5% [97.8, 99.0%]	<0.001	*
Independent	3.3% [2.7, 3.4%]	90.3% [87.8, 92.3%]	0.001	*	0.2% [-0.1, 0.4%]	98.3% [97.8, 98.7%]	0.247	
Third Party's Lab	1.0% [0.1, 1.4%]	92.6% [90.5, 94.3%]	0.289		0.9% [0.9, 0.9%]	97.6% [96.9, 98.2%]	0.000	*
NRL	-2.2% [-2.3, -1.8%]	95.8% [92.9, 97.5%]	0.207	*	4.2% [2.7, 6.4%]	94.4% [91.3, 96.4%]	<0.001	*
US FDA	-2.2% [-3.6, -1.3%]	95.8% [94.2, 97.0%]	0.038	*	0.4% [0.4, 0.4%]	98.1% [97.3, 98.6%]	0.047	*
FIND Diagnostic	18.6% [14.6, 22.8%]	75.0% [67.8, 81.1%]	<0.001	*	0.9% [0.6, 1.3%]	97.6% [96.4, 98.4%]	0.008	*
Netherland CIDC	-0.2% [-0.3, 0.0%]	93.8% [90.5, 96.0%]	0.825		0.5% [0.4, 0.7%]	98.0% [97.0, 98.7%]	0.060	
Doherty	2.7% [1.5, 4.5%]	90.9% [86.1, 94.1%]	0.055		0.8% [0.4, 1.5%]	97.7% [96.2, 98.6%]	0.037	*

SeroTracker

INTERPRETATION & RECOMMENDATIONS

Substantial heterogeneity in assays and assay evaluation techniques, which could bias seroprevalence estimates by up to ± 9.5%.

What should future serosurveys do?

- (1) Seroprevalence studies should consider adopting third-party or independently evaluated assays
- (2) Assay properties are context-specific. Important to validate these assays using the sorts of samples they will be deployed on, to maximize the correspondence between the context they were evaluated in and the context they were deployed in (i.e., spectrum bias)
- (3) Statistical test adjustments should employ validated assay performance data

SeroTracker

LESSONS LEARNED IN THE STUDY EXECUTION

Manufacturers do not evaluate assay performance against a testing sample with diverse composition.

Quantitative assays have become more common, but are still not prevalent.

Unless you become a client purchasing the product, not all manufacturers report assay features

Details of serological assays are often incomplete in serosurvey studies.

- The combined performance of multiple testing algorithms was rarely reported.
- Self-developed assays were self-certified and were used in 18.2% studies, the performance of which cannot be cross-validated.

SeroTracker

LIMITATIONS

- Focused on seroprevalence studies:
  - Excluded studies done exclusively in confirmed COVID-19 cases and vaccinated individuals
  - Population-based contexts may not translate entirely on the patient or clinical level
- Extracted overall Sn./Sp. or Sn./Sp. from the farthest time point available post-symptom onset (often ~30-60 days)
- Could not evaluate self-developed assays due to insufficient data

SeroTracker

ACKNOWLEDGEMENTS

We would like to thank all members of the SeroTracker team, who built the foundation for this study by maintaining an up-to-date database of seroprevalence studies, and especially Xiaomeng Ma for leading the project from conceptualization through manuscript preparation.

SeroTracker

Q & A