Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
An evaluation of 4 commercial assays for the detection of SARS-CoV-2 antibodies in a predominantly mildly symptomatic low prevalence Australian population

Michael C. Wehrhahn, Suzanne J. Brown, James P. Newcombe, Smathi Chong, Jenny Evans, Melanie Figtree, Laurence Hainke, Linda Hueston, Sadid Khan, Elizabeth Marland, Matthew V.N. O’Sullivan, Helen Powell, Jhumur Roy, Lynette Waring, Megan Yu, Jennifer Robson

A total of 1080 individual patient samples (158 positive serology samples from confirmed, predominantly mildly symptomatic COVID-19 patients and 922 serology negative including 496 collected pre-COVID) from four states in Australia were analysed on four commercial SARS-CoV-2 serological assays targeting antibodies to different antigens (Roche Elecsys and Abbott Architect: nucleocapsid; Diasorin Liaison and Euroimmun: spike). A subset was compared to immunofluorescent antibody (IFA) and micro-neutralisation. Sensitivity and specificity of the Roche (n = 1033), Abbott (n = 806), Diasorin (n = 1034) and Euroimmun (n = 175) were 93.7% / 99.5%, 90.2% / 99.4%, 88.6% / 98.6% and 91.3% / 98.8%, respectively. ROC analysis with specificity held at 99% increased the sensitivity for the Roche and Abbott assays from 93.7% to 98.7% (cut-off 0.21) and 90.2% to 94.0% (cut-off 0.91), respectively. Overall seropositivity of samples increased from a maximum of 23% for samples 0-7 days post-onset of symptoms (dpos), to 61% from samples 8-14dpos and 93% from those >14dpos. IFA and microneutralisation values correlated best with assays targeting antibodies to spike protein with values >80 AU/mL on the Diasorin assay associated with neutralising antibody. Detectable antibody was present in 22/23 (96%), 20/23 (87%), 15/23 (65%) and 9/22 (41%) patients with samples >180dpos on the Roche, Diasorin, Abbott and microneutralisation assays respectively. Given the low prevalence in this community, two-step algorithms on initial positive results saw an increase in the positive predictive value (PPV) of positive samples (39%–65% to ≥98%) for all combinations. Similarly accuracy increased from a range of 98.5%–99.4% to ≥99.8% assuming a 1% seroprevalence. Negative predictive value (NPV) was high (≥99.8%) regardless of which assay was used initially.

1. Introduction

As of December 15, 2020, more than 73 million cases of COVID-19 have been diagnosed causing over 1.6 million deaths worldwide.

While diagnosis has relied largely on SARS-CoV-2 nucleic acid amplification testing (NAAT), for a small number of patients with equivocal or negative NAAT results, serology has been instrumental in clarifying the true infection status of a case. Serology has a greater role in
retrospectively diagnosing COVID-19 especially in asymptomatic cases [1,2]. This in turn improves estimates of attack rate, case fatality rate and reproduction number (R0) in a population [3–5]. Serology may also have prognostic value, with antibody titres found to correlate with severity of infection [4,6–9], and can assist public health investigations of outbreaks [5].

In the longer term, studies may be able to assess whether herd immunity against SARS-CoV-2 has been achieved [3]. If high-throughput commercial Enzyme Linked Immunosorbent Assays (ELISA) or Chemiluminescent Microparticle Immunoassays (CMIA/CLIA) correlate with neutralising antibody (Nab) titres, an immune population may be determined who are lower risk for returning to frontline work [3,5]. At present, serological correlates of immunity post-vaccination are yet to be determined and published studies of Nab following infection with SARS-CoV-2 remain limited [6,10–14].

Commercial assays generally target one of two proteins: the spike or the nucleocapsid protein with uncertainty about which targets are the most sensitive or correlate best with Nab [15,16].

While published studies to date have compared commercial assays head-to-head [17–21] most have been in higher prevalence areas. Australia and New Zealand have been fortunate in largely eliminating transmission in the community but this has presented a unique diagnostic challenge to understand the relative performance of commercial test platforms for use in a very low incidence community. We sought to investigate the reliability of four commercial assays for use within our laboratory network in Australia.

2. Methodology

Commercial assay testing (Roche Elecsys Anti-SARS-CoV-2 assay targeting IgM/IgA/IgG to nucleocapsid protein, Abbott Architect SARS-CoV-2 IgG targeting nucleocapsid protein, Diasorin Liaison SARS-CoV-2 S1/S2 IgG targeting the S1 and S2 domains of the spike protein, Euroimmun Anti-SARS-CoV-2 IgG targeting the S1 domain of the spike protein) was performed at 4 laboratories across 4 states of Australia according to the assays’ instructions for use (IFU). Due to limited availability of the Euroimmun assay and residual stored sera, samples from COVID-19 patients were prioritised over specificity samples for testing using this assay. The Institute of Clinical Pathology and Medical Research, NSW performed an in-house immunofluorescent antibody (IFA) assay [22] and a microneutralisation assay [23] as previously described. All data was de-identified and verbal consent of COVID-19 patients was obtained and approval for this study obtained from the Sullivan Nicolaides Pathology Low Risk Ethics Committee.

Sensitivity analysis was performed on stored sera from confirmed patients with COVID-19 diagnosed by NAAT as defined by local guidelines and also household contacts seropositive for IgG by IFA in the absence of NAAT being performed [24,25]. The majority of cases were diagnosed by the Seegene Allplex 2019-nCoV Assay (targeting E, N and RdRp genes) or an in-house developed Taqman assay targeting the E and N genes.

Only the latest sample for each patient was selected. Seropositive samples from confirmed COVID-19 cases had to be positive by at least IFA if tested or in the absence of IFA, at least one commercial assay. A two-step algorithm analysing the different possible combinations of commercial assays to optimise positive predictive value (PPV) was evaluated [26].

The COVID-19 cases and three cohorts used to assess the specificity of the assays under evaluation are depicted in Table 1.

A novel illness severity score (ISS) assessing 19 symptoms and need for hospitalisation, intensive care support and mechanical ventilation (HIM) was developed and evaluated on a subset of patients as depicted in Table 2. Duration of symptoms (DS) was also assessed and capped at 30 days.

Equivocal results for Euroimmun and Diasorin were considered positive for statistical analysis. Sub-analysis was performed with

| Table 1 | Sensitivity and Specificity samples. |
|---------|-------------------------------------|
| Sensitivity cohort | n (%) | Specificity cohort | n (%) |
| Qld 77 (45) | Cohort 1 pre-COVID | 411 |
| NSW 71 (41) | -antenatal | 324 |
| Victoria 14 (8) | -healthy adults 2019 | 57 |
| WA 11 (6) | -children 2019 | 30 |
| Qld 67 (42) | Cohort 2* | 378 |
| NSW 66 (42) | Cross-reactivity (pre-COVID) | 235 |
| Victoria 14 (9) | Cross-reactivity (COVID) | 143 |
| WA 11 (7) | | 38 |
| Total COVID-19 cases | 173 (16) | Total specificity samples | 907 |
| Total seropositive (% of total overall) | 158 (15) | pre-COVID | 646 |
| | | | 71 |
| | | Total overall samples 1080 |
| NSW New South Wales, Qld Queensland, WA Western Australia. ARI: Acute respiratory infection. |

Including but not limited to patients diagnosed with the following respiratory pathogens: influenza A/B n = 55, parainfluenza n = 16, respiratory syncytial virus n = 12, adenovirus n = 10, rhinovirus n = 3, Mycoplasma pneumoniae n = 16, Bordetella pertussis n = 14, Chlamydia pneumoniae n = 10, Legionella longbeachae n = 8; non-respiratory pathogens positive for EBV n = 53, CMV n = 43, Parvovirus B19 n = 19, Hepatitis A/B/C n = 26, HIV 5, Ross River virus n = 17; or highly elevated autoantibodies: Rheumatoid factor n = 19, Antinuclear antibody n = 28.

serology values obtained from IFA, ‘nucleocapsid assays’ (NA), ie Abbott and Roche assays and ‘spike assays’ (SA), ie Diasorin and Euroimmun.

Receiver operator characteristic (ROC) curve analysis was undertaken to determine the area under the curve (AUC), and sensitivity, specificity and accuracy, along with 95 % confidence intervals (CI), were calculated. Assay cut-off values were optimised by interrogation of ROC results.

Pairwise comparisons of sensitivity and specificity between assays were made. As many of the same samples were assessed by each assay, the diagnostic measures were correlated with each other, and to account for this a procedure using 2000 bootstrap replicates was adopted [27]. Sensitivity was tested while holding specificity at 99 %, then specificity for this a procedure using 2000 bootstrap replicates was adopted [27].

We evaluated the relationship between IFA and assay values using linear regression analysis, with adjustment for the covariates age and sex. As the IFA distribution was skewed, a normalising square root transformation was applied before subsequent analysis. Associations with 0.05/9 = 0.0056. The same associations were also examined in sex-stratified analyses.

Data were analysed in the R statistical computing environment, version 4.0.0, including the package pROC [27].

3. Results

There were 173 patients with confirmed COVID-19 by NAAT (n = 169) or IFA (n = 4) who had serology samples evaluated. Of these, 90
(52 %) were male and the mean age of patients was 50.7 years (range 7–85 years). Severity of disease data for 163 patients indicated that 5 (3 %) were asymptomatic, 131 (80 %) were symptomatic but non-hospitalised, 18 (11 %) were hospitalised/non-ICU patients and 9 (6 %) were ICU patients. Thirty day mortality was 0 %.

Of the 173 individual samples from confirmed cases and 907 negative samples, there were 158/1080 (14.6 %) seropositive patients. Time since illness onset available for 147 patients found 139/147 (95 %) were ≥10 days-post-symptom onset (dpos). Sensitivity, specificity and accuracy results are presented in Table 3. In the subset of 175 samples common to all assays, there were no differences in AUC among the 4 assays (all p > 0.042), nor in specificity when sensitivity was held at 95 % (all p > 0.15). When specificity was held at 99 %, there was a significant difference in sensitivity only between the Roche and Diasorin assays (p = 0.003).

When assay cutoff values for the Roche and Abbott assays were optimised (1.0 to 0.21 and 1.4 to 0.91, respectively) to achieve 99 % specificity, sensitivity increased for both assays (Table 4).

Twenty-nine samples from 22/158 (14 %) COVID-19 patients were found to be discrepant on ≥1 commercial assays (Table 5). For a family of three COVID-19 cases, one had a positive result only on SA (16–32 dpos), one only on NA (12 dpos) and one with positive results from at least one of SA and NA (17dpos). False negative results were found in 9/29 (31 %), 14/29 (48 %), 14/29 (48 %), and 17/29 (59 %) Euroimmun, Roche, Abbott and Diasorin results respectively. Twenty-nine samples from 22/158 (14 %) COVID-19 patients were found to be discrepant on ≥1 commercial assays (Table 5). For a family of three COVID-19 cases, one had a positive result only on SA (16–32 dpos), one only on NA (12 dpos) and one with positive results from at least one of SA and NA (17dpos). False negative results were found in 9/29 (31 %), 14/29 (48 %), 14/29 (48 %), and 17/29 (59 %) Euroimmun, Roche, Abbott and Diasorin results respectively. Two families were only positive by IFA but negative/equivocal on commercial assays (although values for these two patients were above the mean negative value of pre-COVID samples).

Sensitivities by time period are presented in Table 6. Seroconversion occurred by at least 12 dpos on commercial assays (range 5–12dpos). Samples remained positive beyond 180 days in the majority of patients by the Roche (96 %), Diasorin (87 %) and Abbott assays (65 %) (Fig. 1).

The mean cutoff value from the 3 specificity cohorts ranged between 0.08–0.09, 0.06–0.11, 4.3–4.5 and 0.3 (positive cutoff of 1.0, 1.4, 15.0 and 1.1) for Roche, Abbott, Diasorin and Euroimmun assays respectively. Commercial assays were also compared with 51 positive and 47 negative IFA results with high positive agreement (Diasorin 43/51 (84 %), Abbott 45/51 (88 %), Euroimmun 46/51 (90 %) and Roche 48/51 (94 %)) and negative agreement (Diasorin and Roche 47/47 (100 %) and Abbott and Euroimmun 46/47 (98 %)). From regression analysis after

### Table 3
Diagnostic performance of 4 assays at the published cut-offs for detection of COVID-19 disease (a) in all 1080 samples, of which 158 (14.6 %) tested positive; and (b) in subset of 175 subjects with measurements on all 4 assays, of which 92 (52.6 %) tested positive. Values are expressed as percentages (95 % CI), and as a decimal fraction for AUC.

|                  | Roche | Abbott | Diasorin | Euroimmun |
|------------------|-------|--------|----------|------------|
| (a) All samples  |       |        |          |            |
| N                | 1033  | 806    | 1034     | 175        |
| Cut-off          | 1     | 1.4    | 12*      | 0.8*       |
| AUC              | 0.997 (0.994, 1) | 0.994 (0.986, 1) | 0.977 (0.961, 1) | 0.990 (0.980, 1) |
| Sensitivity      | 93.7 (89.9, 95.7) | 90.2 (85.0, 95.5) | 88.6 (83.5, 93.0) | 91.3 (84.8, 96.7) |
| Specificity      | 99.5 (99.1, 99.9) | 99.4 (98.8, 99.9) | 98.6 (97.8, 99.3) | 98.8 (96.4, 100) |
| Accuracy         | 98.6 (98.0, 98.3) | 97.9 (96.9, 98.8) | 97.1 (96.1, 98.1) | 94.9 (91.4, 97.7) |
| (b) Subset of 175, with values on all 4 assays |       |        |          |            |
| N                | 175   | 175    | 175      | 175        |
| Cut-off          | 1     | 1.4    | 12*      | 0.8*       |
| AUC              | 0.999 (0.996, 1) | 0.989 (0.976, 1) | 0.984 (0.967, 1) | 0.990 (0.980, 1) |
| Sensitivity      | 93.5 (88.0, 97.8) | 89.1 (82.6, 94.6) | 87.0 (79.4, 93.5) | 91.3 (84.8, 96.7) |
| Specificity      | 100 (100, 100) | 97.6 (94.0, 100) | 97.6 (94.0, 100) | 98.8 (96.4, 100) |
| Accuracy         | 96.6 (93.7, 98.9) | 93.1 (89.1, 96.6) | 92.0 (87.4, 95.4) | 94.9 (91.4, 97.7) |

* Cut-offs presented for Diasorin and Euroimmun are the equivocal and borderline ranges respectively; positive for Diasorin and Euroimmun is considered ≥15 and ≥1.1 respectively.

### Table 4
Diagnostic performance of 4 assays for detection of COVID-19 disease, when specificity is set at 99 %. Values are expressed as percentages (95 % CI).

|                  | Roche | Abbott | Diasorin | Euroimmun |
|------------------|-------|--------|----------|------------|
| N                | 1033  | 806    | 1034     | 175        |
| Published Cut-off| 1     | 1.4    | 12*      | 0.8*       |
| 99 % Specificity | 0.210 | 0.91   | 16.7     | 1.55       |

* Cut-offs presented for Diasorin and Euroimmun are the equivocal and borderline ranges respectively; positive for Diasorin and Euroimmun is considered ≥15 and ≥1.1 respectively.
adjustment for age, there was a stronger association between IFA and SA values (Diasorin: $R^2 = 59\%$; Euroimmun: $R^2 = 61\%$;) than between IFA and NA values (Roche: $R^2 = 10\%$; Abbott: $R^2 = 23\%$).

Thirty microneutralisation results from 27 patients were available (range 32–203 dpos). Fourteen patients (52\%) had detectable Nab (6/6 samples collected $<180$ dpos, 9/22 (41\%) samples $\geq 180$ dpos).

Significantly higher titres were found in those who were hospitalised, while samples collected $\geq 180$ dpos had lower titres ($p < 0.003$). A minority of patients (1/14, by Roche; 2/14 by Diasorin and 3/14 by Abbott) were negative by all commercial assays but had detectable Nab. Increasing assay values only on the SA correlated with presence of Nab.

All 8 patients with an assay value $>80$ AU/mL by Diasorin were found to have detectable Nab with higher values in males than females (Fig. 2).

Eighteen possible two-step algorithms (including optimised sensitivities for the Roche and Abbott assays) were analysed assuming a population prevalence of 1\% based on recent data indicating low seroprevalence in our community [28]. PPV after the second assay improved from a range of 0.39–0.65 to $\geq 0.98$ for all combinations. The highest PPV was obtained from the combination of the Roche and Abbott assays (0.996). Accuracy increased from a range of 0.985–0.994 to $\geq 0.998$. The most accurate combination was the optimised Roche assay followed by the Euroimmun assay (0.999). Calculations for the Roche followed by Diasorin combination are illustrated in Fig. 3.

Twenty pairs of household contacts (HHC) (spouses/partners/family members) providing 54 samples collected within 4 dpos of each other, with range of collection: 7–219 dpos, were matched with 18 pairs of unrelated patients (URP). No significant difference was found between the HHC and URP within-pair assay differences for IFA or the 4 commercial assays, even when this was limited to the subset of patients with probable same exposure (Table 7).

The constructed ISS, evaluated in 62 seropositive patients, ranged from 0 to 17/30 with a mean value of 9. The presence of HIM was more commonly demonstrated in males than females (13/35 vs 2/27). Mean DS was 20 days (range 1 to $\geq 30$ days). ISS was not found to be significantly associated with any serology outcomes. DS showed a significant positive association only with SA ($p = 0.001$) and HIM was significantly positively associated only with IFA and SA (both $p < 0.001$). There was a divergence between male and female groups for Diasorin, Euroimmun,

---

**Table 5**

Discrepant results in 22 COVID-19 positive patients.

| Patient | dpos | Roche | Abbott | Euroimmun | Diasorin | Diasorin |
|---------|------|-------|--------|-----------|----------|----------|
| D3      | 39   | 12.3  | 3.25   | 1.35      | 5.86     | N        |
| D6      | 16   | 1.4   | 3.73   | 0.62      | N        | 4.79     |
| D22     | 35   | 3.81  | 1.52   | 0.93      | EQ       | 6.33     |
| D38     | 42   | 5.07  | 1.01   | 0.64      | N        | 6.44     |
| D29     | 41   | 3.97  | 2.07   | 0.76      | N        | 4.63     |
| D33     | 40   | 0.223 | 0.53   | 3.1       | P        | 43.4     |
| D36     | 36   | 59.9  | 6.55   | 1.4       | P        | 9.32     |
| D42     | 41   | 1.23  | 0.66   | 0.99      | EQ       | 7.57     |
| D50     | 43   | 1.59  | 0.77   | 0.99      | EQ       | 7.62     |
| D70     | 39   | 0.0872| 0.03   | N         | 0.94     | EQ       | 3.8      |
| D71     | 48   | 23.9  | 3.03   | 0.34      | N        | 71.5     |
| D72     | 57   | 23.5  | 3.32   | 0.5       | N        | 25.2     |
| S12     | 16   | 0.166 | N      | 0.6       | N        | 2.262    |
| S12     | 25   | 0.228 | N      | 0.73      | N        | 4.282    |
| S12     | 32   | 0.393 | N      | 0.64      | N        | 4.409    |
| S13     | 17   | 0.792 | N      | 1.86      | P        | 1.705    |
| S14     | 12   | 1.7   | 2.42   | 0.6665    | N        | 3.96     |
| S49     | 13   | 1.23  | 0.83   | 0.5651    | N        | 5.98     |
| S49     | 22   | 0.39  | 0.85   | 1.207     | N        | 9.07     |
| S49     | 29   | 0.613 | N      | 0.3       | N        | 1.489    |
| S55     | 17   | 0.495 | N      | 1.29      | N        | 2.337    |
| S54     | 19   | 0.644 | N      | 2.2       | P        | 2.694    |
| S57     | 23   | 0.672 | N      | 2.34      | P        | 0.7569   |
| S57     | 31   | 2.55  | N      | 3.01      | 0.9117   | EQ       | 8.42     |
| S57     | 40   | 7.15  | N      | 3.49      | 0.8868   | EQ       | 9.14     |
| S62     | 33   | 6.79  | N      | 3.57      | N        | 1.909    |
| S63     | 15   | 0.88  | N      | 0.35      | N        | 0.9439   |
| S63     | 27   | 0.752 | N      | 0.42      | N        | 2.57     |
| S72     | 26   | 0.665 | N      | 3.27      | N        | 0.7775   |

**Table 6**

Sensitivity of various assays for detection of SARS-CoV-2 antibodies by increasing time intervals; All Positive: positive by at least one assay; dpos: days post onset of symptoms.

| dpos | n (%) | All Positive n (%) | Roche n (%) | Abbott n (%) | Euroimmun n (%) | Diasorin n (%) | IFA n (%) |
|------|-------|-------------------|------------|-------------|-----------------|----------------|-----------|
| 1–7  | 32 (13)| 12/32 (38)        | 3/32 (9)   | 4/27 (15)   | 5/22 (23)       | 6/32 (19)      | 1 (%)     |
| >2–8 | 14    | 190 (78)          | 187/190 (99)| 171/190 (90)| 154/182 (85)    | 103/111 (93)   | 171/190 (90) |
| >14  | 28    | 148 (60)          | 147/148 (99)| 139/148 (94)| 122/144 (85)    | 69/74 (93)     | 133/148 (90) |
| >60  | 40 (16)| 40/40(100)        | 39/40 (98) | 28/40 (70)  | 36/40 (90)      | 22/25 (88)     | 72/75 (96) |
| >180 | 25 (9) | 25/25(100)        | 22/23 (96) | 15/23 (65)  | 20/23 (87)      | 20/23(87)      |           |
| Total| 245   | 214/245 (87)      | 186/245 (77)| 168/228 (74)| 113/149 (76)    | 107/245 (76)   |           |

M.C. Wehrhahn et al.
In females, age was found to be significantly positively associated with IFA and all assay values (all $p \leq 0.004$); while HIM, DS and ISS were not. By contrast, in males, age was not associated with any outcome (all $p > 0.05$); HIM showed positive associations only with SA and IFA (both $p \leq 0.004$); and DS and ISS were not significantly associated with any assay outcomes.

Mean/median ratios of SA:NA values for Diasorin:Roche, Diasorin:Abbott, Euroimmun:Roche and Euroimmun:Abbott were significantly higher in hospitalised compared with non-hospitalised patients and highest in those in ICU (Fig. 5).

4. Discussion

This evaluation found satisfactory performance and comparability of
specificity was >98.5 % (range 98.6–99.5 %), and sensitivity/accuracy was high (range 88.6–93.7 %/94.9–98.6 % respectively). Sensitivity/specificity varied by 0.2–1.1 %/0.5–1.8% respectively when analysis was confined to samples that were measured by all 4 assays. The area under the curve (AUC) was >97 % for all assays, indicating high discriminatory power. When analysed by time period, sensitivity increased from a maximum of 61 %<14 dpos to 93 %>14 dpos and 98 %>60 dpos. The positive cases in our dataset experienced predominantly mild to moderate COVID-19 and may explain the reduced sensitivities when compared with those in the assay IFUs.

Discordant results by commercial assays were found in a minority (14 %) of our patients as has been found in other studies [12,17] and may relate to assay design or immune response. Also of note was the difference in persisting seropositivity between the Abbott and Roche assays on samples >180 dpso despite targeting the same nucleocapsid protein in keeping with the indirect 'sandwich assay' design of the Roche assay that preferentially detects more mature antibody. This finding supports the use of more than one assay with differing target antigens or design as has been recommended locally and by others [8,29] to allow serological testing to address a variety of clinical and public health questions.

This is also the only currently available study providing large numbers in the Oceania region where COVID-19 case numbers are relatively limited. Reassuringly, the high specificities of the assays found in this study corresponded closely to those reported by the manufacturers (≤0.5 % difference). We found that initial false positive results can be further reduced by adopting a two-step approach utilising a second assay as found by others [21,30]. Only NA (Roche and Abbott) gained sensitivity by lowering of the positive cutoff without a significant reduction in specificity. A two-step algorithm may also increase the detection of low level antibody only detectable by one type of assay. In one of our laboratories, the use of the Roche assay as the initial screening assay has been enhanced by reflexing any results >0.2 cutoff index to the Diasorin assay. In the other laboratory in our network performing serology, positive results on the Diasorin were reflex tested on the Abbott assay. While these two combinations were not the most accurate or provided the highest PPV, they resulted in the most efficient laboratory workflow using assays targeting two different antigens and resulted in a PPV/accuracy of 98.4 %/99.9 % and 99.0 %/99.8 % respectively. In the near future, with the advent of vaccines that largely target the spike protein, there may be a need to differentiate the development of antibodies and immunity following infection (NA + SA positive) or vaccination (SA positive only) providing further reason for utilising assays that target both antigens. However, careful evaluation of the performance of available assays is required to maximise PPV as was found by Ripperger et al. where their in-house nucleocapsid-specific assay was not suitably discriminatory for use in their two-step algorithm [30].

This is also one of only a few studies that have correlated commercial

### Table 7
Comparison of mean assay difference (diff) between household contact (HHC) pairs and matched unrelated patient pairs (URP); values for continuous variables are presented as mean (standard deviation).

|                  | HHC | HHC (same exposure) | URP | P-value |
|------------------|-----|--------------------|-----|---------|
| Number of patients | 40  | 22                 | 36  |         |
| Number of patient pairs | 20  | 11                 | 18  |         |
| Number of samples | 54  | 30                 | 36  |         |
| % Male           | 48  | 41                 | 47  |         |
| Age mean         | 51.6 (17.1) | 50.9 (17.0) | 51.4 (15.5) | 0.952 0.912 |
| dpos diff mean   | 1.8 (1.6)  | 1.5 (1.4) | 8.0 (6.1)  | 0.005 0.061 |
| Roche pairs      | 27  | 15                 | 18  |         |
| Roche mean diff  | 22.4 (19.6) | 22.4 (19.8) | 32.9 (26.8) | 0.285 0.360 |
| Abbott pairs     | 27  | 15                 | 18  |         |
| Abbott mean diff | 2.3 | 2.9 (2.4) | 2.2 (2.1) | 0.839 0.343 |
| Diasorin pairs   | 27  | 15                 | 18  |         |
| Diasorin mean diff | 52.2 | 50.2 (51.8) | 63.7 (63.1) | 0.616 0.514 |
| Euroimmun pairs  | 17  | 8                  | 9   |         |
| Euroimmun mean diff | 2.6 | 2.8 (2.0) | 2.6 (2.0) | 0.689 0.817 |

![Fig. 3. An example of a 2-step SARS-CoV-2 serology algorithm.](image-url)
assay results with Nab and the only one we are aware of comparing it with reference laboratory IFA results [11]. Overall concordance of negative and positive samples with IFA was high (92–97%) and the high sensitivity of IFA in samples ≥28 dpos (96%) was comparable to that reported recently by members of our group (91% for samples ≥14 dpos) [22]. Nab, currently considered the best potential correlate with protective immunity, was still detectable in 9/22 (41%) of those with samples collected more than 6 months post-illness. Nab also corresponded to IFA results and SA values. Taking into account the higher proportion (67%) of late collected samples (≥180 dpos) of predominantly mildly infected patients in our study, we also found a similar proportion of samples with Diasorin values >80 AU/mL associated with Nab titres of ≥1:160 (5/9, 56% vs 87% in the Diasorin IFU). A negative commercial assay result does not exclude the possibility of detectable Nab and possible protective immunity and caution should be taken when interpreting negative results >6 months following a compatible illness due to the possibility of false negative serology particularly in young females where antibody levels were found to be lowest in both SA and IFA.

There are clear differences in mortality between males and females, despite similar incidence of infection, and this may relate to a differing immune response which is reflected in lower antibody levels [31]. While our novel ISS alone did not correlate with assay values and was likely limited by the retrospective design and associated recall and reporting bias, HIM was found to be significantly associated with higher assay values in the IFA and SA in males. Similar findings were found in one large study where higher antibody levels were seen in older and sicker patients while lower levels were seen in females suggesting gender-specific immune responses towards SARS-CoV-2 [32]. In contrast to other studies that showed lower spike:nucleocapsid antibody ratios in hospitalised and deceased patients [33,34], our study found the higher/highest ratios in the hospitalised/intensive care patients compared with those managed in the community. This may relate to both sampling later in the illness and ultimate recovery of patients in our

Fig. 4. Boxplot of assay values by sex for individual assays and grouped by target antigen (SA and NA) and IFA in the subset of subjects who have an IFA value. 0=female, 1=male.
study, with strong spike-specific antibody responses having been shown to correlate with survival [33].

Since greater inoculum of SARS-CoV-2 has been shown in hamster models to lead to more severe disease [35], by comparing assay values from COVID-19 patients of the same household with matched unrelated patients, we postulated that the household group were more likely to have been exposed to a similar inoculum of SARS-CoV-2 from the same source and therefore might return similar assay values. However, this was not found to be the case and may relate to probable varying inoculum despite the same source/setting in addition to differing host factors amongst household members especially those who were genetically unrelated.

Limitations of this study include its restricted demography relating only to Australian patients and the small sample size of patients with long term results (>180 dpos).

In conclusion, we demonstrate that currently available commercial assays perform well with the majority of samples taken more than 6 months following infection still positive with associated neutralising antibody detected in almost half. In low prevalence areas such as Australia, the use of a two-step algorithm increases the PPV of a positive result while maintaining high NPV/accuracy.

**Funding**

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Transparency document**

The Transparency document associated with this article can be found in the online version.

**Declaration of Competing Interest**

The authors report no declarations of interest.

**Acknowledgements**

The authors would like to thank the patients who consented to providing blood samples, their collectors and laboratory staff involved and to Dr Ian Chambers, Dr Miriam Paul and Dr Rodney Givney for review of the manuscript.

**References**

[1] D. Buitrago-Garcia, D. Egli-Gany, M.J. Cournotte, S. Hossmann, H. Imeri, A. M. Ipekci, G. Salanti, N. Low, Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: a living systematic review and meta-analysis, PLoS Med. 17 (2020), e1003346.

[2] K. Mizumoto, K. Kagaya, A. Zarebski, G. Chowell, Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020, Euro Surveill. (2020) 25.

[3] E.S. Theel, P. Slev, S. Wheeler, M.R. Couturier, S.J. Wong, K. Kadkhoda, The role of antibody testing for SARS-CoV-2: Is there one? J. Clin. Microbiol. (2020) 58.

[4] Q.X. Long, B.Z. Liu, H.J. Deng, G.C. Wu, K. Deng, Y.K. Chen, P. Liao, J.F. Qiu, Y. Lin, X.F. Cai, D.Q. Wang, Y. Hu, J.H. Ren, N. Tang, Y.Y. Xu, Mo Z, Yu LH, F. Gong, X.L. Zhang, W.G. Tian, L. Hu, X.L. Zhang, J.L. Xiang, Liu H.W. Du HX, C. H. Lang, X.H. Liao, S.B. Wu, X.P. Cai, Z. Zhou, M.M. Zhu, J. Wang, C.J. Xue, X.F. Li,
F. Wolff, H. Dahma, C. Duterme, S. Van den Wijngaert, O. Vandenberg, F. Cotton, R. A. Chandrashekar, J. Liu, A. J. Martinot, K. McMahan, N. B. Mercado, L. Peter, L. M. C. Wehrhahn et al.

B. Sun, Y. Feng, X. Mo, P. Zheng, Q. Wang, P. Li, P. Peng, X. Liu, Z. Chen, H. Huang, J. Qu, C. Wu, X. Li, G. Zhang, Z. Jiang, X. Li, Q. Zhu, L. Liu, Profile of A. K. Winter, S. T. Hegde, The important role of serology for COVID-19 control, J. Li, D. C. Zhang, F. Zhang, P. Liu, J. Yuan, Q. Li, T. J. Li, L. S. H. Winstone, C. Kerridge, I. Huettner, J. M. Jimenez-Guarde, J. Campion, M. Chopra, M. Insel, A. P. Capaldi, C. M. Spier, M. Outpatient, L. B. Haagmans, Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) serological assays enable surveillance of low-prevalence communities and reveal durable humoral immunity, Immunity 53 (2020), 925-933.e4.

J. Jin, P. Bai, W. He, F. Wu, X. F. Liu, D. M. Han, S. Liu, J. K. Yang, Gender differences in patients with COVID-19: focus on severity and mortality, Front. Public Health 8 (2020) 152.