Title

Cross-platform comparison of highly-sensitive immunoassays for inflammatory markers in a COVID-19 cohort[1]

Running title: COMPARISON OF INFLAMMATORY MARKER ASSAYS IN COVID-19 COHORT

Koji Abe[*‡], Joanne C. Beer[†‡], Tran Nguyen[*], Ishara S. Ariyapala[†], Tyson H. Holmes[*], Wei Feng[†], Bingqing Zhang[†], Dwight Kuo[†], Yuling Luo[†], Xiao-Jun Ma[†], and Holden T. Maecker[*]

*Institute for Immunity, Transplantation, and Infection,
Stanford University School of Medicine, Stanford, CA 94305
†Alamar Biosciences, Inc., Fremont, CA 94538

‡ Both authors contributed equally to this work.

Corresponding author: Koji Abe
Phone: 650-723-4984
E-mail: kojiabe@stanford.edu

[1]Funded by grant 2U19AI057229 from the National Institutes of Health

Abstract

A variety of commercial platforms are available for the simultaneous detection of multiple cytokines and associated proteins, often employing antibody pairs to capture and detect target proteins. In this study, we comprehensively evaluated the performance of three distinct platforms: the fluorescent bead-based Luminex assay, the proximity extension-based Olink assay, and a novel proximity ligation assay platform known as Alamar NULISAseq. These assessments were conducted on serum samples from the NIH IMPACC study, with a focus on three essential performance metrics: detectability, correlation, and differential expression. Our results reveal several key findings. Firstly, the Alamar platform demonstrated the highest overall detectability, followed by Olink and then Luminex. Secondly, the correlation of protein measurements between the Alamar and Olink platforms tended to be stronger than the correlation of either of these platforms with Luminex. Thirdly, we observed that detectability differences across the platforms often translated to differences in differential expression findings, although high detectability did not guarantee the ability to identify meaningful biological differences. Our study provides valuable insights into the comparative performance of these assays, enhancing our understanding of their strengths and limitations when assessing complex biological samples, as exemplified by the sera from this COVID-19 cohort.

Keywords: multiplexed immunoassay, cytokines, PLA, PEA, NULISA, COVID-19

45 **Introduction**

46 Multiple commercial platforms offer multiplexed detection of cytokines and related proteins,
47 many using pairs of antibodies to capture and detect target proteins. In Luminex assays, the capture
48 antibodies are coated on microbeads, and detector antibodies are labelled with biotin and then streptavidin-
49 phycoerythrin to yield a fluorescent readout (1). In proximity extension and proximity ligation assays
50 (PEA and PLA) (2), pairs of DNA-tagged antibodies bind to the same target protein molecule and facilitate
51 a ligation or polymerase extension reaction whose product can be detected by polymerase chain reaction
52 (PCR) or sequencing. Previous studies have examined the performance of PEA, known as the Olink assay,
53 in comparison to other platforms, including aptamer-based SomaScan (3–6). These comparisons have
54 demonstrated a range of correlations and highlighted various factors that need to be considered when
55 selecting the appropriate platform. Luminex and Olink assays have also been compared (7), with best
56 correlations observed for the more abundant analytes.

57 The NIH IMPACC study (IMmunoPhenotyping of A COVID-19 Cohort) evaluated longitudinal
58 clinical and immunological features of 1164 patients hospitalized for COVID-19 from 2020-2022. The
59 study employed multiple highly multiplexed assays for viral, cellular, protein, and genomic phenotyping
60 (8). Among these was the 92-plex inflammation panel for serum cytokines from Olink Proteomics. We
61 sought to test performance characteristics of this assay against Luminex and a novel PLA platform, Alamar
62 NULISAseq (9). Since the latter comprises over 200 targets, it could provide a broader evaluation of
63 cytokines in the progression of COVID-19 in the IMPACC cohort. Our Luminex comparison assay was a
64 set of three panels totalling 80 target proteins, using kits from Millipore Corporation.

65 For the current evaluation, we chose a small subset of 78 serum samples from 23 IMPACC
66 patients, along with 8 healthy control sera. We performed parallel assays on the Olink, Luminex and
67 Alamar panels, compared the detectability and assessed correlation for common analytes, and compared

68 ability to discriminate COVID-19 from healthy samples, ability to detect changes over time in the COVID-
69 19 patients, and ability to distinguish COVID-19 trajectory groups.
70

Materials and Methods

Cohort

The IMPACC study was conducted to investigate the clinical and immunological characteristics of COVID-19 in hospitalised patients (8, 10). This observational cohort study involved a network of expert clinicians, geneticists, and immunologists with the objective of examining the relationship between the clinical course of COVID-19 and the immune response in racially, ethnically, and geographically diverse adult patient populations across the United States. Serum samples were collected at multiple time points during hospitalization, including enrollment, days 4, 7, 14, 21, and 28. Throughout the hospitalization period, clinical outcome data such as mortality and level of care were diligently documented. For a supplementary analysis, we utilized a framework comprising five distinct disease course trajectories to comprehensively capture the evolution of the clinical course of the disease under investigation. Specifically, trajectory 1 represented a brief length of stay (median 3 days), trajectory 2 denoted an intermediate length of stay (median 7 days), trajectory 3 indicated an intermediate length of stay (median 7 days) with discharge limitations, trajectory 4 corresponded to prolonged hospitalization (median 20 days), and trajectory 5 represented a fatal outcome (8, 11).

Sample preparation

The IMPACC sample subset used in this analysis included 23 COVID-19 patients with samples collected at up to 6 time points each, giving a total of 78 serum samples (mean age 59.3 years, 6 female; see Table I). An additional 8 serum samples were obtained from 8 healthy adult donors (3 female). All samples were stored at -80°C prior to thawing for each assay. These samples were employed without any filtration or modification. This study was approved by the Stanford IRB.

Protein analysis methods

95 Luminex assay (EMD Millipore, Burlington, MA) and Olink immunoassay (Olink Bioscience,
96 Uppsala, Sweden) were performed by the Human Immune Monitoring Center at Stanford University
97 Immunoassay Team. For Luminex assay, Human 80 Plex kits were purchased from EMD Millipore
98 Corporation, Burlington, MA., and run according to the manufacturer's recommendations with
99 modifications described as follows: H80 kits include 3 panels: Panel 1 is Milliplex HCYTA-60K-PX48.
100 Panel 2 is Milliplex HCP2MAG-62K-PX23. Panel 3 includes the Milliplex HSP1MAG-63K-06 and
101 HADCYMAG-61K-03 (Resistin, Leptin and HGF) to generate a 9 plex. PBS was added to wells for
102 reading in the Luminex FlexMap3D Instrument with a lower bound of 50 beads per sample per cytokine.
103 Custom Assay Chex control beads were purchased and added to all wells (Radix BioSolutions,
104 Georgetown, Texas).

105 Assay Chex beads were used for non-specific binding correction in the Luminex assay. The
106 procedure encompassed log-transforming the MFI values of both target proteins and CHEX4 beads,
107 followed by mean-centering the log-transformed CHEX4 MFI. Subsequently, a linear regression was
108 performed, regressing protein MFI values against centered CHEX4 MFI values. The resulting non-specific
109 binding-corrected values were derived by adding the regression residuals to the intercept estimate.

110 For Olink immunoassay, the samples were subjected to Olink inflammatory panel multiplex assay
111 (Target 96 Inflammation), according to the manufacturer's instructions. The inflammatory panel includes
112 92 proteins associated with human inflammatory conditions. Briefly, an incubation master mix containing
113 pairs of oligonucleotide-labeled antibodies to each protein was added to the samples and incubated for 16
114 hours at 4 ºC. Each protein was targeted with two different epitope-specific antibodies increasing the
115 specificity of the assay. Presence of the target protein in the sample brought the partner probes in close
116 proximity, allowing the formation of a double strand oligonucleotide polymerase chain reaction (PCR)
117 target. On the following day, the extension master mix in the sample initiated the specific target sequences
118 to be detected and generated amplicons using PCR in 96 well plate. For the detection of the specific

119 protein, Dynamic Array Integrated Fluidic Circuit (IFC) 96x96 chip was primed, loaded with 92 protein
120 specific primers and mixed with sample amplicons including three inter-plate controls (IPS) and three
121 negative controls (NC). Real time microfluidic qPCR was performed in Biomark (Fluidigm, San
122 Francisco, CA) for the target protein quantification. Data were analyzed using real time PCR analysis
123 software via ΔΔCt method and Normalized Protein Expression (NPX) manager. One unit NPX difference
124 amounts to the doubling of the protein concentration.

125 Alamar's NULISAseq 200-plex Inflammation panel targets mostly inflammation and immune
126 response-related cytokines and chemokines (9). First, serum samples were centrifuged at 10,000g for 10
127 min. Supernatant from the serum samples were then analyzed using Alamar's NULISAseq proteomic
128 platform. For NULISA, the capture antibody is conjugated with partially double-stranded DNA containing
129 a poly-A tail and a target-specific barcode, whereas the detection antibody is conjugated with another
130 partially double-stranded DNA containing a biotin group and a matching target-specific barcode. When
131 both antibodies are incubated with a sample containing the target molecule, an immunocomplex is formed.
132 The formed immunocomplexes are captured by added paramagnetic oligo-dT beads and subsequent dT-
133 polyA hybridization, and the sample matrix and unbound detection antibodies are removed by washing.
134 As dT-polyA binding is sensitive to salt concentration, the formed immunocomplexes are then released
135 into a low-salt buffer. After removing the dT beads, a second set of paramagnetic beads coated with
136 streptavidin is introduced to capture the immunocomplexes in the solid phase a second time. Subsequent
137 washes remove free unbound capture antibodies, resulting in essentially pure immunocomplexes on the
138 beads. Then, a ligation mix containing T4 DNA ligase and a specific DNA ligator sequence is added,
139 allowing ligation of the proximal ends of DNA attached to the paired antibodies and thus generating a new
140 DNA reporter molecule containing unique target-specific barcodes. The levels of the DNA reporter are
141 then quantified by Next Generation Sequencing. Data normalization was done using an internal control
142 spiked into each sample well to remove potential technical variation introduced during data collection.

143 Then data was rescaled and log2-transformed to yield NULISA Protein Quantification (NPQ) units for
144 downstream statistical analyses.
145 Each sample was measured singlicate in each assay platform.
146
147 *Statistical analysis*
148 Data analysis utilized log2 normalized signal intensity across platforms. All observations,
149 including those falling below limit of detection (LOD) or outside of the dynamic range for each assay,
150 were included in the subsequent statistical analyses. The number of common protein targets across
151 platforms were as follows: 33 (across three platforms), 56 (Alamar NULISAseq and Olink Target 96), 58
152 (Alamar NULISAseq and Luminex H80), and 38 (Olink Target 96 and Luminex H80) (Supplemental Data
153 A). All statistical analyses were conducted using R (version 4.3.0) (12).
154
155 *Detectability*
156 To determine detectability, LODs recommended by the respective manufacturers were employed.
157 For all three platforms, LOD is defined as the mean plus 3 times the standard deviation of the blank
158 samples. Detectability was defined as the percentage of measured values out of the 78 total samples above
159 the LOD for each protein target. Since blank samples were not included in the Luminex assay run, LODs
160 were obtained for an independent Luminex run using the same reagent lots.
161
162 *Correlation*
163 Since association between platforms are not always linear, nonparameteric Spearman and Kendall
164 correlations were used to assess correlations across platforms using baseline (first-visit) COVID-19
165 samples. Additionally, to accommodate repeated measures in the entire longitudinal dataset, the R package
166 'rmcorr' (13, 14) was utilized to investigate the common within-participant correlation. The rmcorr method

is based on a form of ANCOVA and assumes linearity and errors that are independent, identically and normally distributed.

Differential expression: COVID-19 patients at enrollment versus healthy controls

Differential expression for COVID-19 patients at enrollment (n=23) versus healthy controls (n=8) was assessed using a linear regression model with disease status as predictor of target expression. For each inter-platform comparison, Benjamini-Hochberg false discovery rate (FDR) correction (15) was applied to the relevant subset of targets and significance was set at 5% FDR. Pearson correlation of the estimated coefficients (i.e., log2(fold change)) was calculated for shared targets for each platform pair.

Differential expression: Change over time in COVID-19 patients

Change in target expression over time in the COVID-19 group (n=23) was assessed using linear mixed effects models with categorical time fixed effects accounting for the 6 time points and a random subject intercept. The joint significance of the time fixed effect was tested using a likelihood ratio test (LRT), as implemented in R package 'lmerTest' (16). For each inter-platform comparison, FDR correction was applied to the relevant subset of targets and significance was set at 5% FDR. Pearson correlation of the LRT chi-squared statistics was calculated for shared targets for each platform pair.

Logistic regression: Predicting COVID-19 outcome using protein expression levels at enrollment

An exploratory analysis was carried out to assess whether protein expression levels at enrollment could predict the most severe (trajectory groups 4 and 5; n=14) versus the least severe (trajectory groups 1, 2, and 3; n=9) COVID-19 patient outcomes. Logistic regression models were fit for each target in each panel using target expression as predictor of the binary trajectory group outcome. Results were evaluated using unadjusted and FDR-adjusted p-values at a 5% significance level.

191

192 **Results**

193 *Overall detectability*

194 Overall across-target mean detectability (i.e. percent of signal values above LOD) for the platforms
195 was 94.7% for Alamar NULISAseq (203 targets), 86.5% for Olink Target 96 (92 targets), and 84.1% for
196 Luminex H80 (80 targets).

197

198 *Detectability with 33 common targets in 3 assays*

199 The detectability of protein targets varied considerably across the three assays for the 33 shared
200 targets (Figures 1A, 1C). In the case of Alamar NULISAseq and Luminex H80, only 1 of 33 protein targets
201 (CCL3 and IL20, respectively) exhibited ≤ 50% detectability. Olink Target 96 had 7 targets with ≤ 50%
202 detectability (IL13, IL20, IL33, IL4, IL5, LIF, TSLP), with five proteins (IL13, IL20, IL4, IL5, and TSLP)
203 consistently falling below the LOD in at least 75% of the samples. For these 33 targets, mean
204 detectabilities were 96.1% for Alamar, 83.5% for Olink, and 87.2% for Luminex.

205

206 *Detectability in pairwise cross-platform comparisons for common targets*

207 Alamar NULISAseq and Olink Target 96 (Figure 1B, left): Of 56 common targets, Alamar showed
208 one target with ≤ 50% of samples above LOD (CCL3), while Olink showed 9 (IL13, IL20, IL2RB, IL33,
209 IL4, IL5, LIF, NGF, TSLP). Alamar mean detectability was 97.3%; Olink mean was 86.5%.

210 Alamar NULISAseq and Luminex H80 (Figure 1B, middle left): Of 58 common targets, Alamar
211 NULISAseq showed one target with ≤ 50% of samples above LOD (CCL3), while Luminex showed 4
212 (FASLG, IFNA2, IL20, IL22). Alamar mean detectability was 96.4%; Luminex mean was 85.6%.

213 Olink Target 96 and Luminex H80 (Figure 1B, middle right): Of 38 common targets, Olink showed
214 9 targets with ≤ 50% of samples above LOD (IL13, IL1A, IL2, IL20, IL33, IL4, IL5, LIF, TSLP), while

Luminex showed 4 (IL1A, IL2, IL20, LTA). Olink mean detectability was 80.5%; Luminex mean was 83.7%.

In summary, in pairwise cross-platform detectability comparisons for common targets, Alamar showed the highest detectability, followed by Luminex and then Olink. The non-specific binding correction improved detectability for Luminex (Supplemental Figures 1A, 1B). Without the correction, Luminex detectability was below Olink in all comparisons.

Correlation with 33 common targets in 3 assays

We compared Kendall's coefficient of concordance for 33 common targets using baseline COVID samples (Figure 2A). Alamar and Olink showed moderate to high level of concordance (Kendall's coefficient mean: 0.56, median: 0.71, range: -0.14 to 0.91). Alamar and Luminex showed a moderate to low level of concordance (Kendall's coefficient mean: 0.48, median: 0.41, range: -0.30 to 0.85), which was lower compared to Alamar and Olink. Olink and Luminex showed a low level of concordance (Kendall's coefficient mean: 0.36, median: 0.42, range: -0.16 to 0.82), which was lower compared to both Alamar and Olink, and Alamar and Luminex. Overall, the results suggested varying levels of concordance between different measurement methods. Alamar and Olink exhibited the highest concordance, followed by Alamar and Luminex, and then Olink and Luminex.

Correlation in pairwise cross-platform comparisons for common targets

Figure 2B displays the distribution of Kendall's coefficient in each platform comparison for all of the shared targets. Specifically, the comparison between Alamar and Olink (56 targets) indicated a moderate to high level of concordance (Kendall's coefficient mean: 0.56, median: 0.70, range: -0.15 to 0.93; Figure 2B, left). The Alamar and Luminex comparison (58 targets) suggested a moderate to low level of concordance (Kendall's coefficient mean: 0.36, median: 0.40, range: -0.30 to 0.85; Figure 2B,

239 middle), which was lower compared to Alamar and Olink. Moreover, findings indicated a similar level of
240 concordance between Olink and Luminex (38 targets; Kendall's coefficient mean: 0.35, median: 0.40,
241 range: -0.16 to 0.82; Figure 2B, right) when compared to Alamar and Luminex. Taken together, these
242 results highlighted varying levels of concordance between different measurement methods, mirroring the
243 observation of the highest concordance by Alamar and Olink in comparison involving 33 shared targets.
244
245 *Association between detectability and correlation*
246 The association between the difference in detectability and the correlation is depicted in Figure
247 2C. Notably, in each cross-platform comparison, strong correlations were observed when the differences
248 in detectability were minimal. Conversely, larger disparities in detectability generally corresponded to
249 lower correlation. Furthermore, Alamar NULISAseq exhibited greater detectability than Olink Target 96
250 and Luminex H80 for targets exhibiting lower overall correlation (more targets to the right of the dotted
251 vertical line at zero in Figures 2C left and middle).
252
253 *Differential expression in COVID-19 patients versus healthy controls, 33 common targets in 3 assays*
254 For the 33 common targets, 6 were differentially expressed in all 3 panels (Figure 3A and 3B). All
255 of these targets were upregulated (CCL8, CXCL10, HGF, IL10, IL18, IL6). Alamar uniquely identified
256 upregulated CX3CL1 and LIF, and downregulated IL4. Olink found 9 unique differentially expressed
257 targets, all of which were downregulated (CCL4, CXCL5, IL12B, IL13, IL17A, IL20, IL5, KITLG,
258 TSLP). However, several of these (IL13, IL20, IL5, TSLP) also showed ≤ 50% detectability in Olink
259 (Figure 1C). This might be attributed to a potential batch effect, as 7 of 8 control samples were run
260 separately from the COVID-19 samples, and the control sample run had higher LODs for most targets
261 versus the COVID-19 sample run (Supplemental Figure 2). The Luminex panel did not uniquely identify

262 any targets. However, when non-specific binding correction was not applied, Luminex uniquely identified
263 IL7 and TNF (Supplemental Figure 1C).
264
265 *Differential expression in COVID-19 patients versus healthy controls, pairwise comparisons for common*
266 *targets*
267 Alamar NULISAseq and Olink Target 96: Of 56 common targets, Alamar and Olink identified 17
268 common significant targets (Figure 3C, left). Alamar and Olink log2(fold change) estimates for the 56
269 targets were positively correlated (r = 0.83). Alamar identified an additional 6 upregulated (CSF1,
270 CX3CL1, FGF21, IL15RA, IL17C, LIF) and 2 downregulated (IL4, TGFB1) targets, and Olink identified
271 11 additional targets, all of which were downregulated (CXCL5, FGF19, IL12B, IL13, IL20, IL2RB, IL5,
272 KITLG, TNFSF11, TNFSF12, TSLP). Again, several of these (IL2RB, IL20, IL13, IL5, TSLP) showed ≤
273 50% detectability in Olink, and their significance may be related to a batch effect.
274 Alamar NULISAseq and Luminex H80: Of 58 common targets, Alamar and Luminex identified
275 11 common significant targets (Figure 3C, middle). Alamar and Luminex log2(fold change) estimates for
276 the 58 targets were positively correlated (r = 0.75). Alamar identified an additional 9 upregulated (CCL7,
277 CSF3, CX3CL1, CXCL1, IFNA2, IL16, IL1RN, LIF, VCAM1) and 5 downregulated (CCL13, FASLG,
278 FLT3LG, IL4, TNFSF10) targets, and Luminex identified no additional targets.
279 Olink Target 96 and Luminex H80: Of 38 common targets, Olink and Luminex identified 7
280 common significant targets (Figure 3C, right). Olink and Luminex log2(fold change) estimates for the 38
281 targets were positively correlated (r = 0.72). Olink identified an additional 2 upregulated (CCL7, CXCL1)
282 and 16 downregulated (CCL11, CCL13, CCL4, CXCL5, FLT3LG, IL12B, IL13, IL17A, IL1A, IL2, IL20,
283 IL5, KITLG, LTA, TNFSF10, TSLP) targets, and Luminex identified 1 additional upregulated target
284 (CSF1). Some significant Olink targets showed ≤ 50% detectability (IL13, IL1A, IL2, IL20, IL5, TSLP).

285 In summary, in pairwise across-platform comparisons of differential expression in COVID-19 at
286 enrollment versus healthy control, Pearson correlation of log2(fold change) estimates were reasonably
287 high and fairly similar, ranging from the highest for Alamar-Olink (r = 0.83) to lowest for Olink-Luminex
288 (r = 0.72). All platform pairs had some findings in common, and each platform showed some unique
289 findings, with Luminex showing the fewest unique findings. In some cases, unique findings corresponded
290 with superior detectability, e.g., Alamar IL4 and LIF (which had ≤ 50% detectability for Olink), and
291 Alamar FASLG and IFNA2 (which had ≤ 50% detectability for Luminex)).
292
293 *Differential expression in COVID-19 patients versus healthy controls, full panel results*
294 The Alamar panel detected 84 (41.4%) of 203 targets that were differentially expressed (73
295 upregulated, 11 downregulated). The Olink panel detected 41 (44.6%) of 92 targets that were differentially
296 expressed (17 upregulated, 24 downregulated). The Luminex panel detected 12 (15%) of 80 targets that
297 were differentially expressed (10 upregulated, 2 downregulated). Thus the Alamar and Olink panels
298 yielded similar rates of differentially expressed targets, while Luminex showed the smallest, and Olink
299 had a larger percentage of downregulated targets. Supplemental Figure 3A shows volcano plots of full-
300 panel differential expression results for each platform.
301
302 *Differential expression: Change over time in COVID-19, 33 common targets in 3 assays*
303 For the 33 common targets, 8 showed significant changes over time in COVID-19 patients in all 3
304 panels (Figure 4A and 4B). Luminex uniquely identified 2 additional targets (IL7, VEGFA); Alamar
305 identified 5 additional targets (CSF1, IL13, IL20, IL4, TNF), and Olink identified 2 (CCL3, KITLG). The
306 non-specific binding correction resulted in 2 more significant targets for Luminex (CCL13, TNFSF10;
307 Supplemental Figure 1C). Figure 5 shows expression over time for selected targets. Supplemental Figure
308 4 shows all remaining targets that were significant in one or more platforms.

Differential expression: Change over time in COVID-19, pairwise comparisons for common targets

Alamar NULISAseq and Olink Target 96: Of 56 shared targets, Alamar and Olink identified 17 common significant targets (Figure 4C, left). Alamar and Olink chi-squared statistics for the 56 targets were positively correlated (r = 0.66). Alamar identified 9 additional targets (CSF1, FGF19, FGF23, IL13, IL20, IL2RB, IL4, IL5, TNF), and Olink identified 3 additional targets (CCL3, IL18R1, KITLG). Notably, IL13, IL20, IL2RB, IL4, and IL5 had ≤ 50% detectability for Olink, and CCL3 had ≤ 50% detectability for Alamar.

Alamar NULISAseq and Luminex H80: Of 58 common targets, Alamar and Luminex identified 11 common significant targets (Figure 4C, middle). Alamar and Luminex chi-squared statistics for the 58 targets were positively correlated (r = 0.52). Alamar identified 13 additional targets (CCL7, CSF1, FASLG, FLT3LG, IFNA2, IL12p70, IL13, IL17A, IL20, IL4, THPO, TNF, TNFSF10), and Luminex identified 2 additional targets (IL1RN, VEGFA). Notably, IL20 and FASLG had ≤ 50% detectability for Luminex versus 100% for Alamar; IL17A had 53.4% detectability for Luminex versus 94.2% for Alamar; IL12p70 had 58.1% detectability for Luminex versus 87.2% for Alamar.

Olink Target 96 and Luminex H80: Of 38 common targets, Olink and Luminex identified 7 common significant targets (Figure 4C, right). Olink and Luminex chi-squared statistics for the 38 targets were positively correlated (r = 0.58). Olink identified 7 additional targets (CCL3, CCL7, FLT3LG, IL17A, KITLG, LTA, TNFSF10), and Luminex identified 2 additional targets (IL5, VEGFA). Notably, IL17A had higher detectability for Olink (83.7%) than Luminex (53.5%), and IL5 had higher detectability for Luminex (97.7%) than Olink (20.9%).

In summary, in pairwise across-platform comparisons of differential expression over time in COVID-19 patients, correlation of LRT chi-squared statistics was highest for Alamar-Olink (r=0.66), and these two platforms had the largest number of common significant findings (17 targets). Luminex

333 identified the least number of significant targets in the pairwise comparisons, and several targets that
334 Luminex missed also had low detectability for Luminex. In each comparison, Alamar identified more
335 significant changes over time than the other platform.
336
337 *Differential expression: Change over time in COVID-19, full panel results*
338 Considering the full panels for each platform, the Alamar panel detected a total of 86 (42.4%) of
339 203 targets showing significant change over time, the Olink panel detected a total of 29 (31.5%) of 92
340 targets showing significant change over time, and the Luminex panel detected a total of 15 (18.8%) of 80
341 targets showing significant change over time. Supplemental Figure 3B shows volcano plots of full-panel
342 differential expression results for each platform.
343
344 *Logistic regression: Predicting COVID-19 outcome using protein expression levels at enrollment*
345 Due to small sample size (9 least severe versus 14 most severe), none of the panels showed
346 significant FDR-adjusted results. Based on unadjusted $p < 0.05$ significance, Alamar (NULISAseq)
347 showed 1 target with odds ratio significantly below one (CCL3) and 6 targets with odds ratio significantly
348 above one (GDF15, IL12p70, IL17B, IL35, MDK, S100A9) for the most versus least severe trajectory
349 COVID-19 patients. However, CCL3 and IL35 had low detectability for Alamar (32.6% and 5.8%,
350 respectively). For the other targets, area under the curve (AUC) ranged from 0.833 to 0.746. Olink (Target
351 96) and Luminex (H80) showed no significant targets. However, for Luminex without non-specific
352 binding correction, THPO odds ratio was significantly above 1, and VCAM1 had a large effect size (odds
353 ratio=1470, p=0.06, AUC=0.803). Supplemental Figure 3C shows volcano plots and ROC curves.
354 Detailed results, including detectability, inter-assay correlation, and differential expression results
355 for each assay and target can be found in Supplemental Data A.
356

357 **Discussion**

358 In this study, we conducted the first cross-platform comparison to evaluate the measurement of
359 inflammatory markers in COVID-19 samples using three advanced and highly sensitive technologies:
360 bead-based immunoassay Luminex H80, proximity extension assay (PEA) Olink Target 96, and a novel
361 proximity ligation assay (PLA), Alamar NULISAseq. The primary objectives of this study were twofold:
362 firstly, to provide insights into platform selection for future investigations; and secondly, to enhance the
363 interpretation of previous research findings by establishing a comprehensive understanding of the
364 immunoassays employed for measuring the inflammatory-related proteome in the context of COVID-19.

365 The Alamar, Olink, and Luminex panels are complementary in that they each include a unique set
366 of markers. This study highlights strengths and weaknesses of the overall panels and for common marker
367 subsets. The observed differences in detectability between platforms illustrate assay-specific variations in
368 sensitivity, emphasising the importance of considering this assay characteristic when evaluating protein
369 abundance. We found that overall detectability for the platforms was highest for Alamar, followed by
370 Olink and then Luminex.

371 Larger differences in detectability generally corresponded to lower correlation of protein measures
372 across platforms. However, there were some exceptions where detectability was high in both platforms
373 and yet correlation was low. For example, CSF1 showed 100% detectability by both Alamar and Olink,
374 but Kendall coefficient was 0.05. In general, Alamar and Olink correlation tended to be greater than
375 correlation of either platform with Luminex.

376 In many cases, detectability differences also translated to differences in differential expression
377 findings. For example, Alamar showed the highest detectability for IL13 and IL20, and was the only
378 platform able to detect significant increases in these targets over time in COVID-19 patients. Alamar and
379 Luminex both showed higher detectability than Olink for IL5, and thus were able to show significant
380 increase in IL5 over time. Increasing trends for IL13 and IL5 have been previously reported in severe

381 COVID-19 (17), and IL13 has been implicated as a driver of COVID-19 severity (18). Similarly, the
382 higher detectability of Alamar (92.9%) for IFNA2 than Luminex (40.0%) enabled Alamar to demonstrate
383 statistically significant changes for IFNA2 in baseline and longitudinal comparisons, as expected for an
384 antiviral response (19).

385 However, high detectability did not always correspond to significant findings. For example,
386 Alamar alone was able to detect differences at baseline and increase over time in IL4, which had low
387 detectability in Olink (24.4%), but 97.7% detectability in Luminex. IL4 increases have also been noted
388 previously in severe COVID-19 (17). Furthermore, Luminex did not show changes in FLT3LG even with
389 detectability 98.8%, while both Alamar (60.9% detectability) and Olink (100% detectability) showed
390 significant FLT3LG baseline differences and changes over time in COVID-19. This shows that high
391 detectability does not guarantee the ability to detect meaningful biological differences.

392 Low inter-platform correlation for targets with high detectability and incongruent differential
393 expression findings for these targets might result from differences in dynamic range between platforms.
394 Even when samples are above LOD, a smaller dynamic range will limit the ability to detect true biological
395 variation. Additionally, incongruent findings could be due to differences in specificity among platforms.
396 While analytical sensitivity relates to the LOD, analytical specificity depends on discrimination between
397 targets and other potentially interfering proteins (20). Non-specific signal could potentially obscure
398 biological differences.

399 The assay platforms studied here rely on some shared mechanisms for ensuring specificity, such
400 as dual antibody recognition. However, they also employ various different strategies to achieve specificity.
401 For Luminex we applied a non-specific binding correction utilising Assay Chex beads (21). Overall
402 Luminex detectability was enhanced through the application of this method (Supplemental Figure 1), as
403 the method estimated and removed non-specific signal on a per-well basis, including the negative control
404 wells used to define LOD. Differential expression analysis after the correction resulted in more findings

405 for some comparisons and fewer for others. Conversely, in the NULISA, the implementation of capture
406 and release steps was designed to yield a significant reduction of non-specific binding.

407 Our study reinforces previous findings in terms of cytokines associated with COVID-19. For
408 example, all 3 platforms showed upregulated IL6, CXCL10, and IL10 at baseline, and these have been
409 identified as an important triad of plasma markers that distinguish more severe COVID-19 patients (22).
410 Furthermore, CCL8, CXCL10, IL18, and IL6 (identified by all 3 platforms), and CCL7 (identified by
411 Alamar and Olink) were found to be consistently upregulated in COVID-19 versus controls in a meta-
412 analysis of five clinical studies (23). Many of the targets identified by the Alamar panel to be predictive
413 of COVID-19 outcomes have been previously associated with outcome severity in COVID-19, e.g.,
414 GDF15 (24, 25), S100A9 (26), IL17B (27), MDK (28), and IL12p70 (29, 30).

415 Cross-platform comparisons showed that Olink identified the most differentially expressed targets
416 at baseline, although for low detectability targets this was likely confounded by a batch effect, followed
417 by Alamar and then Luminex. In longitudinal analyses, Alamar identified the most targets, followed by
418 Olink and then Luminex.

419 While this study reveals similarities and differences for the 3 assays, it is worth noting some
420 limitations. Despite normalization methods applied to reduce bias, data suggests that some batch effect is
421 still present for the Olink data. Specifically, Olink found more significantly downregulated targets in the
422 baseline COVID-19 versus healthy control comparison relative to the 2 other platforms, and several of
423 these targets showed low detectability. Observed differences in background between the two Olink runs
424 could potentially explain this. In standard practice, it makes sense to exclude such low detectability targets
425 from further analyses, but for the sake of inter-platform comparisons we opted to include them. Also note
426 that while Olink found baseline differences for IL5, IL13, or IL20, Olink did not show longitudinal
427 findings for these while Alamar did, as did Luminex for IL5. In spite of this, many Olink baseline

differential expression significant findings were still in agreement with other platforms. The longitudinal analysis for Olink is not affected by any batch effects, as all COVID-19 samples were run together.

In addition to performance-related criteria such as sensitivity and specificity, other key factors to consider when selecting an immunoassay platform include panel coverage and ease of use. Panel coverage emerges as a pivotal criterion for platform selection. The Alamar panel encompasses over 200 targets, making it a good choice for researchers seeking a wide range of analytes. Olink, while providing 92 targets in its standard offering, presents the potential for comprehensive exploration through the Olink Explore option, which includes more than 3000 targets. Luminex, although starting with fewer targets, stands out for its flexibility and customization capabilities, allowing researchers to run partial plates and to tailor panels according to specific research needs. Thus, the choice of platform can be aligned with the desired scope and specific targets of interest. Ease of use also plays a crucial role in the selection process. The Alamar platform excels in user-friendliness, offering a fully integrated assay process that spans from setup (achievable in under 30 minutes) to the pooling of NGS libraries (completed in less than 6 hours, albeit NGS runs separately). The input volume requirement of 10 µL enhances its accessibility. Olink and Luminex are both more manual and require similar hands-on time to process samples. While Olink has an input volume of only 1 µl, Luminex requires the highest volume at 25 µl. Researchers should weigh the convenience of assay setup and run times based on their specific project requirements.

In this study, we aimed to comprehensively assess the detectability, correlation, and differential expression of protein targets using three different assays: Luminex, Olink, and Alamar. Despite the noted limitations, our findings illuminate aspects of protein quantification within the context of COVID-19, and provide insights into the commonalities and distinctions among these platforms. Specifically, our results reveal that overall detectability for the platforms was highest for the Alamar assay, followed by Olink and Luminex. Additionally, Alamar and Olink showed the greatest concordance in pairwise platform comparisons. Although most of the differences in cross-platform correlation and differential expression

452 results could be explained by differences in sensitivity of different platforms, some might also be
453 attributable to differences in assay dynamic range or specificity. This study thus shows how the diverse
454 characteristics of these platforms have potential implications for the accuracy and robustness of
455 downstream analyses. Only with the combination of high sensitivity, high specificity, and a sufficiently
456 broad dynamic range can biological insights be reliably obtained.
457

Acknowledgements

We thank the IMPACC network and IMPACC patients for access to samples.

461 **References**

462 1. Krishnan, V. V., S. R. Selvan, N. Parameswaran, N. Venkateswaran, P. A. Luciw, and K. S.
463 Venkateswaran. 2018. Proteomic profiles by multiplex microsphere suspension array. *J. Immunol.*
464 *Methods* 461: 1–14.
465 2. Greenwood, C., D. Ruff, S. Kirvell, G. Johnson, H. S. Dhillon, and S. A. Bustin. 2015. Proximity
466 assays for sensitive quantification of proteins. *Biomol. Detect. Quantif.* 4: 10–16.
467 3. Katz, D. H., J. M. Robbins, S. Deng, U. A. Tahir, A. G. Bick, A. Pampana, Z. Yu, D. Ngo, M. D.
468 Benson, Z.-Z. Chen, D. E. Cruz, D. Shen, Y. Gao, C. Bouchard, M. A. Sarzynski, A. Correa, P.
469 Natarajan, J. G. Wilson, and R. E. Gerszten. 2022. Proteomic profiling platforms head to head:
470 Leveraging genetics and clinical traits to compare aptamer-and antibody-based methods. *Sci. Adv.* 8:
471 5164.
472 4. Raffield, L. M., H. Dang, K. A. Pratte, S. Jacobson, L. A. Gillenwater, E. Ampleford, I. Barjaktarevic,
473 P. Basta, C. B. Clish, A. P. Comellas, E. Cornell, J. L. Curtis, C. Doerschuk, P. Durda, C. Emson, C. M.
474 Freeman, X. Guo, A. T. Hastie, G. A. Hawkins, J. Herrera, W. C. Johnson, W. W. Labaki, Y. Liu, B.
475 Masters, M. Miller, V. E. Ortega, G. Papanicolaou, S. Peters, K. D. Taylor, S. S. Rich, J. I. Rotter, P.
476 Auer, A. P. Reiner, R. P. Tracy, D. Ngo, R. E. Gerszten, W. K. O'Neal, and R. P. Bowler. 2020.
477 Comparison of Proteomic Assessment Methods in Multiple Cohort Studies. *Proteomics* 20: 1900278.
478 5. Pietzner, M., E. Wheeler, J. Carrasco-Zanini, N. D. Kerrison, E. Oerton, M. Koprulu, J. Luan, A. D.
479 Hingorani, S. A. Williams, N. J. Wareham, and C. Langenberg. 2021. Synergistic insights into human
480 health from aptamer- and antibody-based proteomic profiling. *Nat. Commun.* 12: 6822.
481 6. Haslam, D. E., J. Li, S. T. Dillon, X. Gu, Y. Cao, O. A. Zeleznik, N. Sasamoto, X. Zhang, A. H.
482 Eliassen, L. Liang, M. J. Stampfer, S. Mora, Z.-Z. Chen, K. L. Terry, R. E. Gerszten, F. B. Hu, A. T.
483 Chan, T. A. Libermann, and S. N. Bhupathiraju. 2022. Stability and reproducibility of proteomic profiles
484 in epidemiological studies: comparing the Olink and SOMAscan platforms. *Proteomics* 22: 2100170.

485 7. Zetlen, H. L., K. T. Cao, K. D. Schichlein, N. Knight, H. T. Maecker, K. C. Nadeau, M. E. Rebuli,
486 and M. B. Rice. 2023. Comparison of multiplexed protein analysis platforms for the detection of
487 biomarkers in the nasal epithelial lining fluid of healthy subjects. *J. Immunol. Methods* 517: 113473.
488 8. Ozonoff, A., J. Schaenman, N. D. Jayavelu, C. E. Milliren, C. S. Calfee, C. B. Cairns, M. Kraft, L. R.
489 Baden, A. C. Shaw, F. Krammer, H. Van Bakel, D. A. Esserman, S. Liu, A. F. Sesma, V. Simon, D. A.
490 Hafler, R. R. Montgomery, S. H. Kleinstein, O. Levy, C. Bime, E. K. Haddad, D. J. Erle, B. Pulendran,
491 K. C. Nadeau, M. M. Davis, C. L. Hough, W. B. Messer, N. I. A. Higuita, J. P. Metcalf, M. A. Atkinson,
492 S. C. Brakenridge, D. Corry, F. Kheradmand, L. I. R. Ehrlich, E. Melamed, G. A. McComsey, R.
493 Sekaly, J. Diray-Arce, B. Peters, A. D. Augustine, E. F. Reed, M. C. Altman, P. M. Becker, N.
494 Rouphael, A. Ozonoff, J. Schaenman, N. D. Jayavelu, C. E. Milliren, C. S. Calfee, C. B. Cairns, M.
495 Kraft, L. R. Baden, A. C. Shaw, F. Krammer, H. Van Bakel, D. A. Esserman, S. Liu, A. F. Sesma, V.
496 Simon, D. A. Hafler, R. R. Montgomery, S. H. Kleinstein, O. Levy, C. Bime, E. K. Haddad, D. J. Erle,
497 B. Pulendran, K. C. Nadeau, M. M. Davis, C. L. Hough, W. B. Messer, N. I. A. Higuita, J. P. Metcalf,
498 M. A. Atkinson, S. C. Brakenridge, D. Corry, F. Kheradmand, L. I. R. Ehrlich, E. Melamed, G. A.
499 McComsey, R. Sekaly, J. Diray-Arce, B. Peters, A. D. Augustine, E. F. Reed, K. McEnaney, B. Barton,
500 C. Lentucci, M. Saluvan, A. C. Chang, A. Hoch, M. Albert, T. Shaheen, A. T. Kho, S. Thomas, J. Chen,
501 M. D. Murphy, M. Cooney, S. Presnell, G. K. Fragiadakis, R. Patel, L. Guan, J. Gygi, S. Pawar, A.
502 Brito, Z. Khalil, C. Maguire, S. Fourati, J. A. Overton, R. Vita, K. Westendorf, R. Salehi-Rad, A.
503 Leligdowicz, M. A. Matthay, J. P. Singer, K. N. Kangelaris, C. M. Hendrickson, M. F. Krummel, C. R.
504 Langelier, P. G. Woodruff, D. L. Powell, J. N. Kim, B. Simmons, I. M. Goonewardene, C. M. Smith, M.
505 Martens, J. Mosier, H. Kimura, A. C. Sherman, S. R. Walsh, N. C. Issa, C. Dela Cruz, S. Farhadian, A.
506 Iwasaki, A. I. Ko, S. Chinthrajah, N. Ahuja, A. J. Rogers, M. Artandi, S. A. R. Siegel, Z. Lu, D. A.
507 Drevets, B. R. Brown, M. L. Anderson, F. W. Guirgis, R. V. Thyagarajan, J. F. Rousseau, D. Wylie, J.
508 Busch, S. Gandhi, T. A. Triplett, G. Yendewa, O. Giddings, E. J. Anderson, A. K. Mehta, J. E.

509 Sevransky, B. Khor, A. Rahman, D. Stadlbauer, J. Dutta, H. Xie, S. Kim-Schulze, A. S. Gonzalez-
510 Reiche, A. Van De Guchte, K. Farrugia, Z. Khan, H. T. Maecker, D. Elashoff, J. Brook, E. Ramires-
511 Sanchez, M. Llamas, A. Rivera, C. Perdomo, D. C. Ward, C. E. Magyar, J. A. Fulcher, Y. Abe-Jones, S.
512 Asthana, A. Beagle, S. Bhide, S. A. Carrillo, S. Chak, G. K. Fragiadakis, R. Ghale, A. Gonzalez, A.
513 Jauregui, N. Jones, T. Lea, D. Lee, R. Lota, J. Milush, V. Nguyen, L. Pierce, P. A. Prasad, A. Rao, B.
514 Samad, C. Shaw, A. Sigman, P. Sinha, A. Ward, A. Willmore, J. Zhan, S. Rashid, N. Rodriguez, K.
515 Tang, L. T. Altamirano, L. Betancourt, C. Curiel, N. Sutter, M. T. Paz, G. Tietje-Ulrich, C. Leroux, J.
516 Connors, M. Bernui, M. A. Kutzler, C. Edwards, E. Lee, E. Lin, B. Croen, N. C. Semenza, B. Rogowski,
517 N. Melnyk, K. Woloszczuk, G. Cusimano, M. R. Bell, S. Furukawa, R. McLin, P. Marrero, J. Sheidy, G.
518 P. Tegos, C. Nagle, N. Mege, K. Ulring, V. Seyfert-Margolis, M. Conway, D. Francisco, A. Molzahn, H.
519 Erickson, C. C. Wilson, R. Schunk, B. Sierra, T. Hughes, K. Smolen, M. Desjardins, S. Van Haren, X.
520 Mitre, J. Cauley, X. Li, A. Tong, B. Evans, C. Montesano, J. H. Licona, J. Krauss, J. B. P. Chang, N.
521 Izaguirre, O. Chaudhary, A. Coppi, J. Fournier, S. Mohanty, M. C. Muenker, A. Nelson, K. Raddassi,
522 M. Rainone, W. E. Ruff, S. Salahuddin, W. L. Schulz, P. Vijayakumar, H. Wang, E. Wunder Jr., H. P.
523 Young, Y. Zhao, M. Saksena, D. Altman, E. Kojic, K. Srivastava, L. Q. Eaker, M. C. Bermúdez-
524 González, K. F. Beach, L. A. Sominsky, A. R. Azad, J. M. Carreño, G. Singh, A. Raskin, J. Tcheou, D.
525 Bielak, H. Kawabata, L. C. Mulder, G. Kleiner, A. S. Lee, E. D. Do, A. Fernandes, M. Manohar, T.
526 Hagan, C. A. Blish, H. N. Din, J. Roque, S. Yang, A. Brunton, P. E. Sullivan, M. Strnad, Z. L. Lyski, F.
527 J. Coulter, J. L. Booth, L. A. Sinko, L. L. Moldawer, B. Borresen, B. Roth-Manning, L.-Z. Song, E.
528 Nelson, M. Lewis-Smith, J. Smith, P. G. Tipan, N. Siles, S. Bazzi, J. Geltman, K. Hurley, G. Gabriele,
529 S. Sieg, T. Vaysman, L. Bristow, L. Hussaini, K. Hellmeister, H. Samaha, A. Cheng, C. Spainhour, E.
530 M. Scherer, B. Johnson, A. Bechnak, C. R. Ciric, L. Hewitt, E. Carter, N. Mcnair, B. Panganiban, C.
531 Huerta, J. Usher, S. P. Ribeiro, M. C. Altman, P. M. Becker, and N. Rouphael. 2022. Phenotypes of
532 disease severity in a cohort of hospitalized COVID-19 patients: Results from the IMPACC study.

533 *eBioMedicine* 83: 104208.

534 9. Wei Feng, Joanne Beer, Qinyu Hao, Ishara S. Ariyapala, Aparna Sahajan, Andrei Komarov, Katie
535 Cha, Mason Moua, Xiaolei Qiu, Xiaomei Xu, Shweta Iyengar, Thu Yoshimura, Rajini Nagaraj, Li
536 Wang, Ming Yu, Kate Engel, Lucas Zhen, Wen Xue, Chen-jung Lee, Chan Ho Park, Cheng Peng,
537 Kaiyuan Zhang, Adrian Grzybowski, Johnnie Hahm, Susanne V. Schmidt, Alexandru Odainic, Jasper
538 Spitzer, Bonn Consortium COVIMMUNE, Kasun Buddika, Dwight Kuo, Lei Fang, Bingqing Zhang,
539 Steve Chen, Eicke Latz, Yiyuan Yin, Yuling Luo, and Xiao-Jun Ma. 2023. NULISA: a novel proteomic
540 liquid biopsy platform with attomolar sensitivity and high multiplexing. *bioRxiv* 2023.04.09.536130.

541 10. Team, I. M. W., on behalf of the I. N. S. Committee, N. Rouphael, H. Maecker, R. R. Montgomery,
542 J. Diray-Arce, S. H. Kleinstein, M. C. Altman, S. E. Bosinger, W. Eckalbar, L. Guan, C. L. Hough, F.
543 Krammer, C. Langelier, O. Levy, K. McEnaney, B. Peters, A. Rahman, J. V. Rajan, S. Sigelman, H.
544 Steen, H. van Bakel, A. Ward, M. R. Wilson, P. Woodruff, C. R. Zamecnik, A. D. Augustine, A.
545 Ozonoff, E. F. Reed, P. M. Becker, N. A. Higuita, M. C. Altman, M. A. Atkinson, L. R. Baden, P. M.
546 Becker, C. Bime, S. C. Brakenridge, C. S. Calfee, C. B. Cairns, D. Corry, M. M. Davis, A. D.
547 Augustine, L. I. R. Ehrlich, E. K. Haddad, D. J. Erle, A. Fernandez-Sesma, D. A. Hafler, C. L. Hough, F.
548 Kheradmand, S. H. Kleinstein, M. Kraft, O. Levy, G. A. McComsey, E. Melamed, W. Messer, J.
549 Metcalf, R. R. Montgomery, K. C. Nadeau, A. Ozonoff, B. Peters, B. Pulendran, E. F. Reed, N.
550 Rouphael, M. Sarwal, J. Schaenman, R. Sekaly, A. C. Shaw, and V. Simon. 2021. Immunophenotyping
551 assessment in a COVID-19 cohort (IMPACC): A prospective longitudinal study. *Sci. Immunol.* 6:
552 eabf3733.

553 11. Diray-Arce, J., S. Fourati, N. Doni Jayavelu, R. Patel, C. Maguire, A. C. Chang, R. Dandekar, J. Qi,
554 B. H. Lee, P. van Zalm, A. Schroeder, E. Chen, A. Konstorum, A. Brito, J. P. Gygi, A. Kho, J. Chen, S.
555 Pawar, A. S. Gonzalez-Reiche, A. Hoch, C. E. Milliren, J. A. Overton, K. Westendorf, J. Abraham, M.
556 Adkisson, M. Albert, L. T. Altamirano, B. Alvarenga, M. L. Anderson, E. J. Anderson, A. Arnett, H.

557 Asashima, M. A. Atkinson, L. R. Baden, B. Barton, K. Beach, E. Beagle, P. M. Becker, M. R. Bell, M.
558 Bernui, C. Bime, A. K. Boddapati, J. L. Booth, B. Borresen, S. C. Brakenridge, L. Bristow, R. Bryant,
559 C. S. Calfee, J. M. Carreño, S. Carrillo, S. Chak, I. Chang, J. Connors, M. Conway, D. B. Corry, D.
560 Cowan, B. Croen, C. S. Dela Cruz, G. Cusimano, L. Eaker, C. Edwards, L. I. R. Ehrlich, D. Elashoff, H.
561 Erickson, D. J. Erle, S. Farhadian, K. Farrugia, B. Fatou, A. Fernandes, A. Fernandez-Sesma, G. K.
562 Fragiadakis, S. Furukawa, J. N. Geltman, R. Ghale, M. C. Bermúdez González, I. M. Goonewardene, E.
563 S. Guerrero, F. W. Guirgis, D. A. Hafler, S. Hamilton, P. Harris, A. N. Hayati, C. M. Hendrickson, N. I.
564 Agudelo Higuita, T. Hodder, S. M. Holland, C. L. Hough, C. Huerta, K. C. Hurley, S. R. Hutton, A.
565 Iwasaki, A. Jauregui, M. Jha, B. Johnson, D. Joyner, K. N. Kangelaris, G. Kelly, Z. Khalil, Z. Khan, F.
566 Kheradmand, J. N. Kim, H. Kimura, A. I. Ko, B. Kohr, M. Kraft, M. Krummel, M. A. Kutzler, J. Lasky-
567 Su, S. Lee, D. Lee, M. Leipold, C. Lentucci, C. Leroux, E. Lin, S. Liu, C. Love, Z. Lu, L. Maliskova, B.
568 R. Manning, M. Manohar, M. Martens, G. A. McComsey, K. McEnaney, R. McLin, E. Melamed, N.
569 Melnyk, K. Mendez, W. B. Messer, J. P. Metcalf, G. Michelotti, E. Mick, S. Mohanty, J. Mosier, L. C.
570 F. Mulder, M. Murphy, K. R. C. Nadeau, E. Nelson, A. Nelson, V. Nguyen, J. Oberhaus, B. Panganiban,
571 K. L. Pellegrini, H. C. Pickering, D. L. Powell, S. Presnell, B. Pulendran, A. H. Rahman, A. S. Rashid,
572 A. Raskin, E. F. Reed, S. P. Ribeiro, A. M. Rivera, J. E. Rogers, A. Rogers, B. Rogowski, R. Rooks, Y.
573 Rosenberg-Hasson, J. Rothman, J. F. Rousseau, R. Salehi-Rad, M. Saluvan, H. Samaha, J. Schaenman,
574 R. Schunk, N. C. Semenza, S. Sen, J. Sevransky, V. Seyfert-Margolis, T. Shaheen, A. C. Shaw, S. Sieg,
575 S. A. R. Siegel, N. Sigal, N. Siles, B. Simmons, V. Simon, G. Singh, L. Sinko, C. M. Smith, K. K.
576 Smolen, L.-Z. Song, K. Srivastava, P. Sullivan, C. Syphurs, J. Tcheou, G. P. Tegos, G. K. Tharp, A.
577 Tong, A. Tsitsiklis, R. F. Ungaro, T. Vaysman, A. Viode, R. Vita, X. Wang, A. Ward, D. C. Ward, A.
578 Willmore, K. Woloszczuk, K. Wong, P. G. Woodruff, L. Xu, S. van Haren, A. van de Guchte, Y. Zhao,
579 C. B. Cairns, N. Rouphael, S. E. Bosinger, S. Kim-Schulze, F. Krammer, L. Rosen, N. D. Grubaugh, H.
580 van Bakel, M. Wilson, J. Rajan, H. Steen, W. Eckalbar, C. Cotsapas, C. R. Langelier, O. Levy, M. C.

581 Altman, H. Maecker, R. R. Montgomery, E. K. Haddad, R. P. Sekaly, D. Esserman, A. Ozonoff, P. M.
582 Becker, A. D. Augustine, L. Guan, B. Peters, and S. H. Kleinstein. 2023. Multi-omic longitudinal study
583 reveals immune correlates of clinical course among hospitalized COVID-19 patients. *Cell Rep. Med.* 4:
584 101079.
585 12. R Core Team (2023). _R: A Language and Environment for Statistical Computing_. R Foundation
586 for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>. .
587 13. Bakdash, J. Z., and L. R. Marusich. 2017. Repeated measures correlation. *Front. Psychol.* 8: 456.
588 14. Bakdash J, Marusich L (2022). _rmcorr: Repeated Measures Correlation_. R package version 0.5.4,
589 <https://CRAN.R-project.org/package=rmcorr>. .
590 15. Benjamini, Y., and Y. Hochberg. 1995. Controlling the False Discovery Rate: A Practical and
591 Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* 57: 289–300.
592 16. Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2017. lmerTest Package: Tests in Linear
593 Mixed Effects Models. *J. Stat. Softw.* 82: 1–26.
594 17. Lucas, C., P. Wong, J. Klein, T. B. R. Castro, J. Silva, M. Sundaram, M. K. Ellingson, T. Mao, J. E.
595 Oh, B. Israelow, T. Takahashi, M. Tokuyama, P. Lu, A. Venkataraman, A. Park, S. Mohanty, H. Wang,
596 A. L. Wyllie, C. B. F. Vogels, R. Earnest, S. Lapidus, I. M. Ott, A. J. Moore, M. C. Muenker, J. B.
597 Fournier, M. Campbell, C. D. Odio, A. Casanovas-Massana, A. Obaid, A. Lu-Culligan, A. Nelson, A.
598 Brito, A. Nunez, A. Martin, A. Watkins, B. Geng, C. Kalinich, C. Harden, C. Todeasa, C. Jensen, D.
599 Kim, D. McDonald, D. Shepard, E. Courchaine, E. B. White, E. Song, E. Silva, E. Kudo, G. DeIuliis, H.
600 Rahming, H.-J. Park, I. Matos, J. Nouws, J. Valdez, J. Fauver, J. Lim, K.-A. Rose, K. Anastasio, K.
601 Brower, L. Glick, L. Sharma, L. Sewanan, L. Knaggs, M. Minasyan, M. Batsu, M. Petrone, M. Kuang,
602 M. Nakahata, M. Campbell, M. Linehan, M. H. Askenase, M. Simonov, M. Smolgovsky, N. Sonnert, N.
603 Naushad, P. Vijayakumar, R. Martinello, R. Datta, R. Handoko, S. Bermejo, S. Prophet, S. Bickerton, S.
604 Velazquez, T. Alpert, T. Rice, W. Khoury-Hanold, X. Peng, Y. Yang, Y. Cao, Y. Strong, R. Herbst, A.

605 C. Shaw, R. Medzhitov, W. L. Schulz, N. D. Grubaugh, C. Dela Cruz, S. Farhadian, A. I. Ko, S. B.
606 Omer, A. Iwasaki, and Yale IMPACT Team. 2020. Longitudinal analyses reveal immunological
607 misfiring in severe COVID-19. *Nature* 584: 463–469.
608 18. Donlan, A. N., T. E. Sutherland, C. Marie, S. Preissner, B. T. Bradley, R. M. Carpenter, J. M.
609 Sturek, J. Z. Ma, G. B. Moreau, J. R. Donowitz, G. A. Buck, M. G. Serrano, S. L. Burgess, M. M.
610 Abhyankar, C. Mura, P. E. Bourne, R. Preissner, M. K. Young, G. R. Lyons, J. J. Loomba, S. J.
611 Ratcliffe, M. D. Poulter, A. J. Mathers, A. J. Day, B. J. Mann, J. E. Allen, and W. A. Petri Jr. 2021. IL-
612 13 is a driver of COVID-19 severity. *JCI Insight* 6: e150107.
613 19. Katze, M. G., Y. He, and M. Gale. 2002. Viruses and interferon: a fight for supremacy. *Nat. Rev.*
614 *Immunol.* 2: 675–687.
615 20. Wilson, R. 2013. Sensitivity and specificity: twin goals of proteomics assays. Can they be
616 combined? *Expert Rev. Proteomics* 10: 135–149.
617 21. Maecker, H. T., Y. Rosenberg-Hasson, K. D. Kolstad, V. D. Steen, and L. S. Chung. 2020. A Novel
618 Utility to Correct for Plate/Batch/Lot and Nonspecific Binding Artifacts in Luminex Data. *J. Immunol.*
619 204: 3425–3433.
620 22. Laing, A. G., A. Lorenc, I. del Molino del Barrio, A. Das, M. Fish, L. Monin, M. Muñoz-Ruiz, D. R.
621 McKenzie, T. S. Hayday, I. Francos-Quijorna, S. Kamdar, M. Joseph, D. Davies, R. Davis, A. Jennings,
622 I. Zlatareva, P. Vantourout, Y. Wu, V. Sofra, F. Cano, M. Greco, E. Theodoridis, J. D. Freedman, S.
623 Gee, J. N. E. Chan, S. Ryan, E. Bugallo-Blanco, P. Peterson, K. Kisand, L. Haljasmägi, L. Chadli, P.
624 Moingeon, L. Martinez, B. Merrick, K. Bisnauthsing, K. Brooks, M. A. A. Ibrahim, J. Mason, F. Lopez
625 Gomez, K. Babalola, S. Abdul-Jawad, J. Cason, C. Mant, J. Seow, C. Graham, K. J. Doores, F. Di Rosa,
626 J. Edgeworth, M. Shankar-Hari, and A. C. Hayday. 2020. A dynamic COVID-19 immune signature
627 includes associations with poor prognosis. *Nat. Med.* 26: 1623–1635.
628 23. Suhre, K., H. Sarwath, R. Engelke, M. U. Sohail, S. J. Cho, W. Whalen, S. Alvarez-Mulett, J.

629 Krumsiek, A. M. K. Choi, and F. Schmidt. 2022. Identification of Robust Protein Associations With
630 COVID-19 Disease Based on Five Clinical Studies. *Front. Immunol.* 12: 781100.
631 24. Gisby, J., C. L. Clarke, N. Medjeral-Thomas, T. H. Malik, A. Papadaki, P. M. Mortimer, N. B.
632 Buang, S. Lewis, M. Pereira, F. Toulza, E. Fagnano, M.-A. Mawhin, E. E. Dutton, L. Tapeng, A. C.
633 Richard, P. D. Kirk, J. Behmoaras, E. Sandhu, S. P. McAdoo, M. F. Prendecki, M. C. Pickering, M.
634 Botto, M. Willicombe, D. C. Thomas, and J. E. Peters. 2021. Longitudinal proteomic profiling of
635 dialysis patients with COVID-19 reveals markers of severity and predictors of death. *eLife* 10: e64827.
636 25. Myhre, P. L., C. Prebensen, H. Strand, R. Røysland, C. M. Jonassen, A. Rangberg, V. Sørensen, S.
637 Søvik, H. Røsjø, M. Svensson, J. Erik Berdal, and T. Omland. 2020. Growth Differentiation Factor 15
638 Provides Prognostic Information Superior to Established Cardiovascular and Inflammatory Biomarkers
639 in Unselected Patients Hospitalized With COVID-19. *Circulation* 142: 2128–2137.
640 26. Chen, L., X. Long, Q. Xu, J. Tan, G. Wang, Y. Cao, J. Wei, H. Luo, H. Zhu, L. Huang, F. Meng, L.
641 Huang, N. Wang, X. Zhou, L. Zhao, X. Chen, Z. Mao, C. Chen, Z. Li, Z. Sun, J. Zhao, D. Wang, G.
642 Huang, W. Wang, and J. Zhou. 2020. Elevated serum levels of S100A8/A9 and HMGB1 at hospital
643 admission are correlated with inferior clinical outcomes in COVID-19 patients. *Cell. Mol. Immunol.* 17:
644 992–994.
645 27. Chamnanphon, M., M. Pongpanich, T. B. Suttichet, W. Jantarabenjakul, P. Torvorapanit, O.
646 Putcharoen, P. Sodsai, C. Phokaew, N. Hirankarn, P. Chariyavilaskul, and V. Shotelersuk. 2022. Host
647 genetic factors of COVID-19 susceptibility and disease severity in a Thai population. *J. Hum. Genet.* 67:
648 295–301.
649 28. Ketenci, S., M. Uygar Kalaycı, B. Dündar, R. Duranay, and A. Şükrü Aynacıoğlu. 2022. Elevated
650 serum midkine levels in severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infected
651 patients. *Int. Immunopharmacol.* 110: 108939.
652 29. Moll-Bernardes, R., A. S. de Sousa, A. V. S. Macedo, R. D. Lopes, N. Vera, L. C. R. Maia, A.

653 Feldman, G. D. A. S. Arruda, M. J. C. Castro, P. M. Pimentel-Coelho, D. C. de Albuquerque, T. C. de
654 Paula, T. A. B. Furquim, V. A. Loures, K. G. D. Giusti, N. M. de Oliveira, F. A. De Luca, M. D. M.
655 Kotsugai, R. A. M. Domiciano, M. F. Santos, O. F. de Souza, F. A. Bozza, R. R. Luiz, and E. Medei.
656 2021. IL-10 and IL-12 (P70) Levels Predict the Risk of Covid-19 Progression in Hypertensive Patients:
657 Insights From the BRACE-CORONA Trial. *Front. Cardiovasc. Med.* 8: 702507.
658 30. Liu, Y., D. Chen, J. Hou, H. Li, D. Cao, M. Guo, Y. Ling, M. Gao, Y. Zhou, Y. Wan, and Z. Zhu.
659 2021. An inter-correlated cytokine network identified at the center of cytokine storm predicted COVID-
660 19 prognosis. *Cytokine* 138: 155365.
661
662
663
664
665

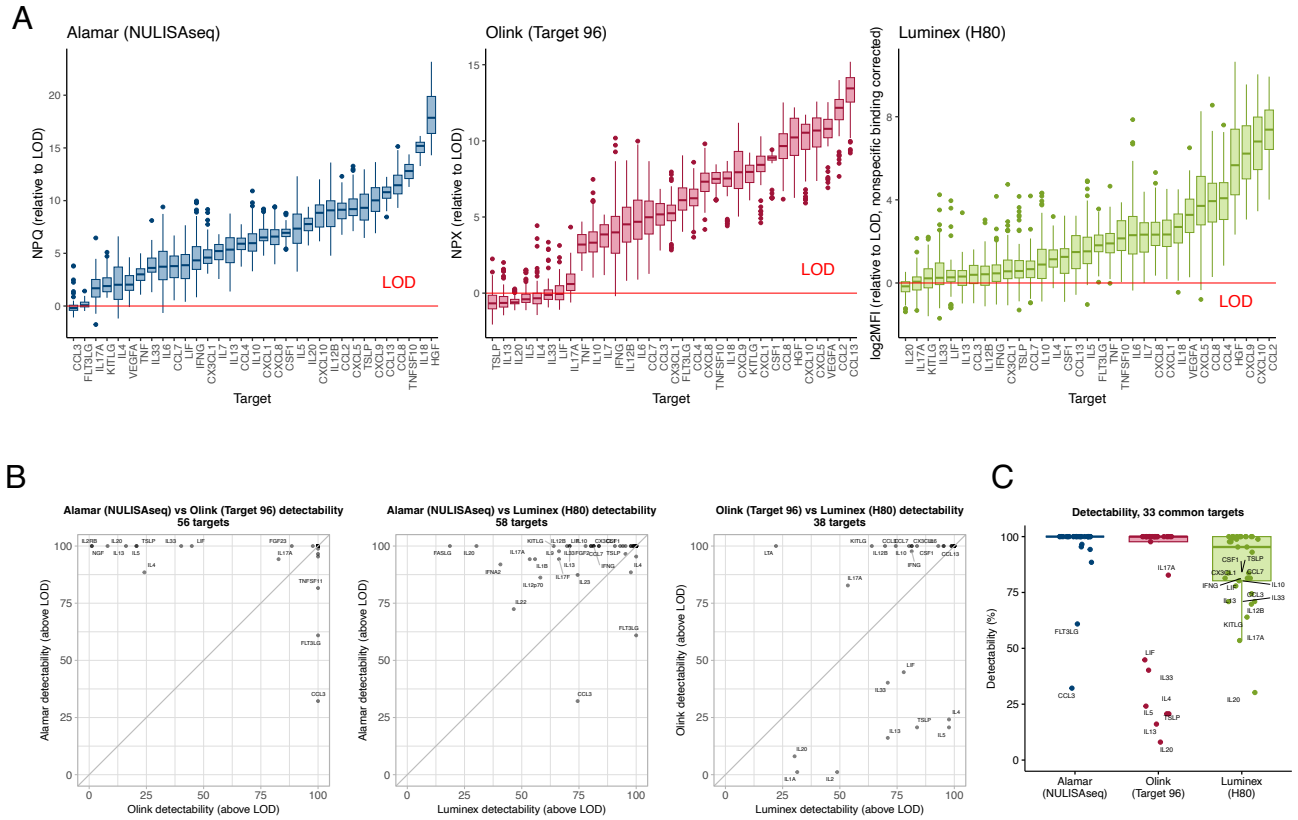

Figure 1. Detectability across platforms. (A) Measured values relative to limit of detection (LOD) for 33 common targets assessed in 78 COVID-19 serum samples from 23 patients and 8 healthy control serum samples. Targets are sorted from left to right based on the median. (B) Comparison of detectability for common targets in each platform pair. (C) Detectability of three platforms for 33 common targets.

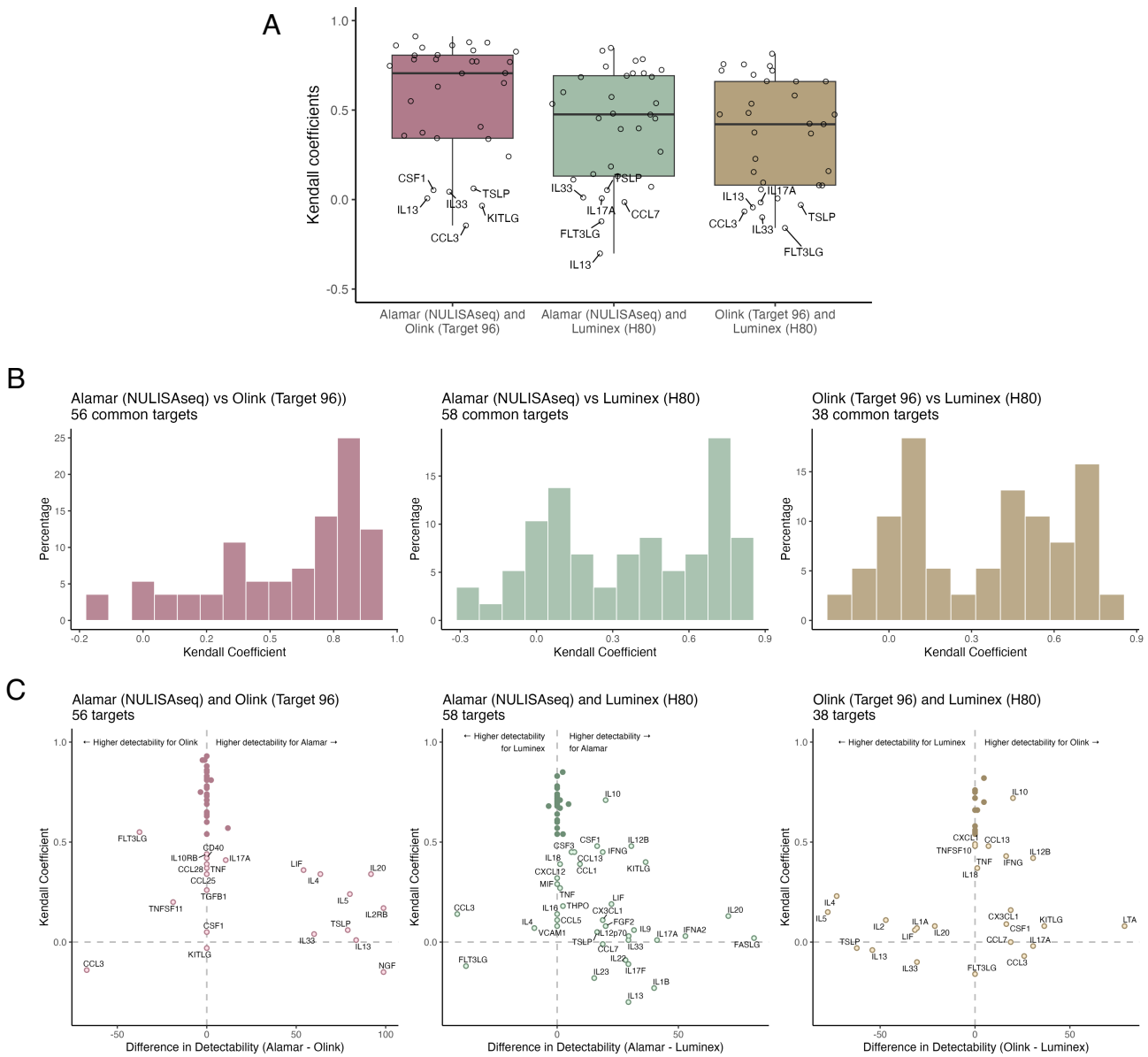

Figure 2. Correlation of measured values across platforms using baseline (Visit 1) samples for COVID-19 (n=23) and Healthy Controls (n=8). (A) Kendall coefficients of each platform pair for 33 common targets. Alamar-Olink showed the highest median correlation, Alamar-Luminex showed second-highest, and Olink-Luminex showed the lowest. (B) Kendall coefficient distributions for all common targets in each platform pair. (C) Association between the difference in detectability (x-axis) and Kendall correlation (y-axis) for common targets in each platform pair using baseline samples. Targets falling on the vertical dashed line showed zero difference in detectability; horizontal dashed line indicates correlation

680 coefficient of zero. Labels highlight protein targets with Kendall coefficient below 0.50 or a difference in
681 detectability greater than 15%. For the most part, larger differences in detectability between platforms
682 corresponds to lower correlation. However, some targets with no detectability difference still showed poor
683 correlation, for example, CSF1 and KITLG for Alamar-Olink, and FLT3LG for Olink-Luminex.
684

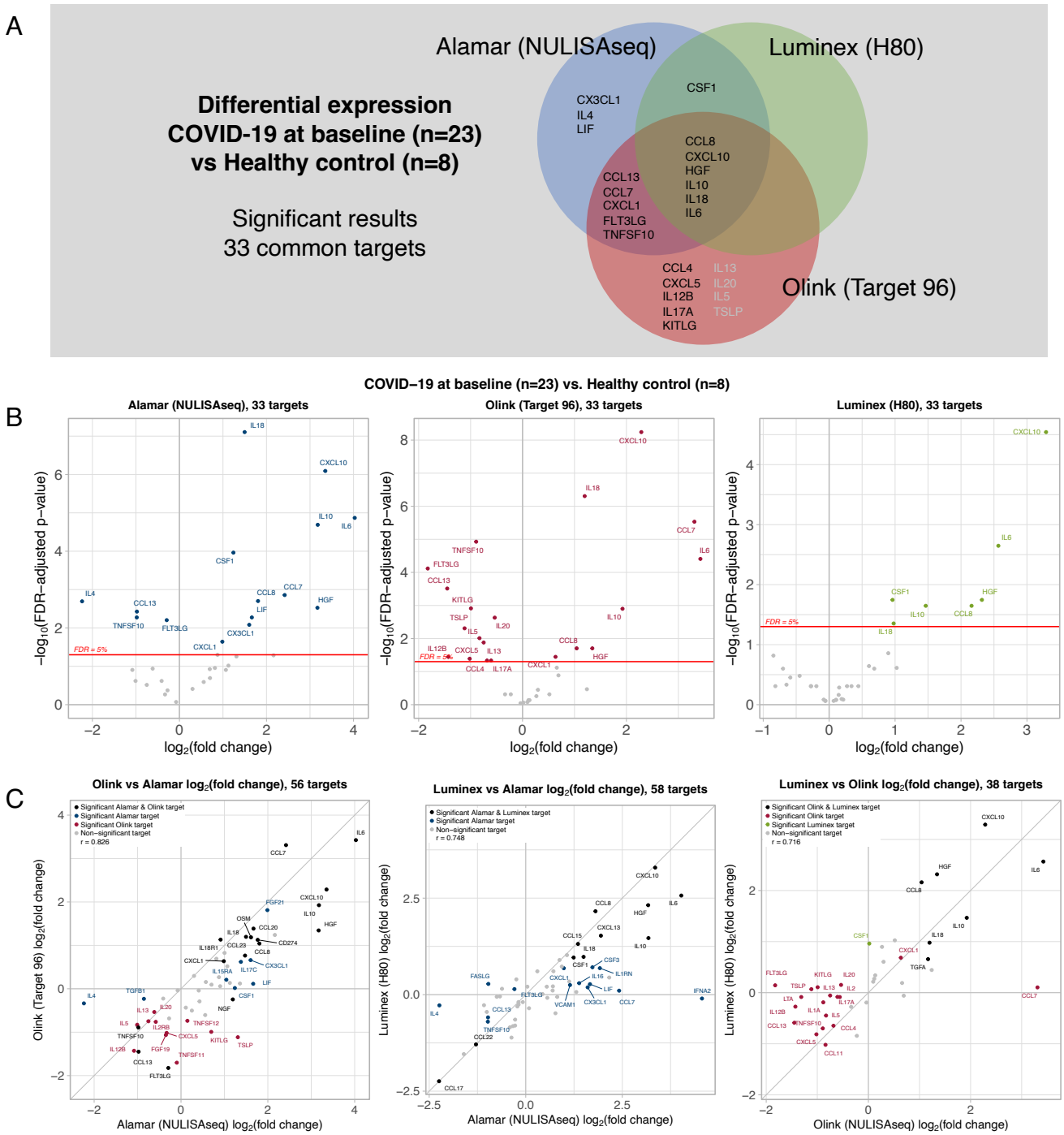

Figure 3. Differential expression comparing inflammatory protein target expression levels for baseline (Visit 1) COVID-19 patients (n=23) and Healthy Controls (n=8). (A) Venn diagram shows overlaps of significant results across panels. Targets in grey text for Olink indicate detectability less than 50%. (B) Volcano plots show -log10 FDR-adjusted p-values versus log2 fold changes for the 33

690 common targets. (C) Scatterplots show the relationship of log2 fold changes for all common targets in

691 each platform pair.

692

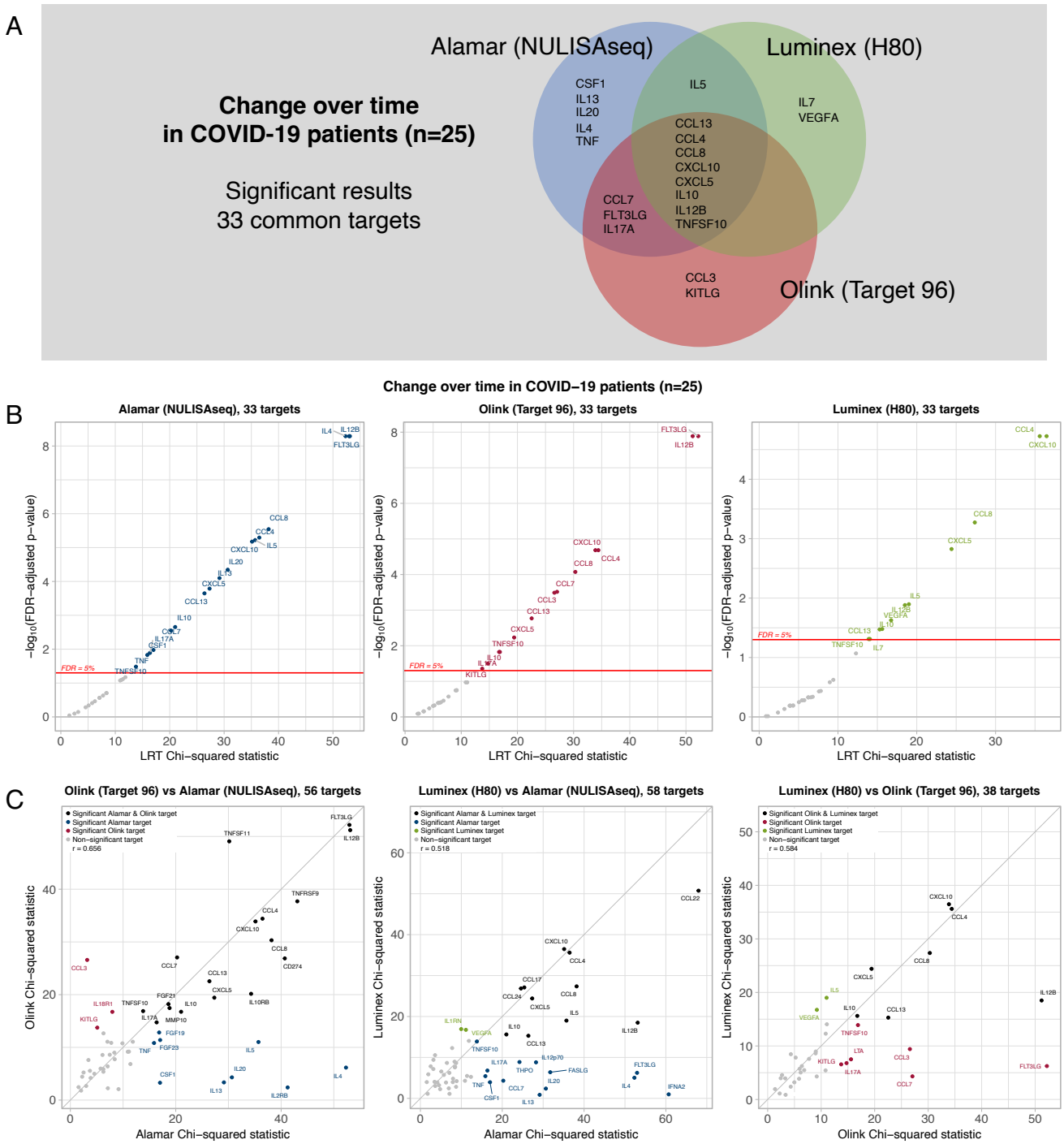

Figure 4. Longitudinal differential expression in COVID patients (n=23). Change over time for inflammatory protein target expression was assessed for the COVID-19 samples using linear mixed effects models with target expression level as outcome, a categorical time fixed effect for each visit, and subject random intercept. The overall significance of the time fixed effect was tested using a likelihood ratio test

698 (LRT). False-discovery rate (FDR) correction was applied to the 33 targets common to all three platforms,
699 and to each subset of overlapping targets for each platform pair. (A) Venn diagram shows overlaps of
700 significant results across platforms. (B) Volcano plots show -log10 FDR-adjusted p-values versus the LRT
701 chi-squared statistic for the 33 common targets. (C) Scatterplots show the relationship of LRT chi-squared
702 statistics for all common targets in each platform pair.
703
704

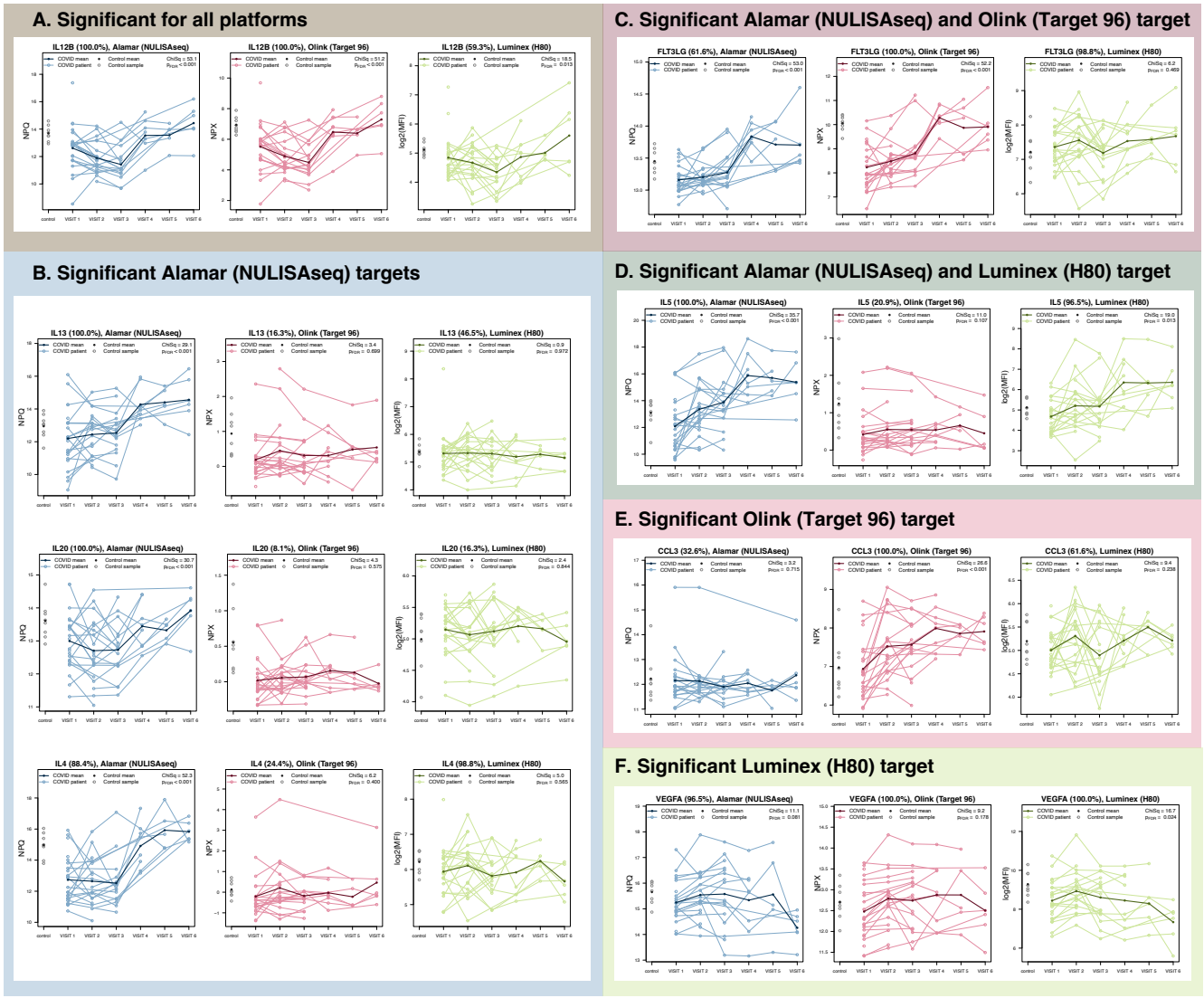

Figure 5. Selected targets showing significant change over time in COVID-19 for one or more platforms. Longitudinal differential expression of inflammatory markers was assessed for 78 serum samples collected from 23 COVID-19 patients over 6 time points. The 6 visits correspond to time of study enrollment (VISIT 1), and days 4, 7, 14, 21, and 28 (VISIT 2-6). Control sample values are shown at the far left side of plots. Values shown in parentheses in plot titles represent target detectability. (A) IL12B was among the targets showing most significant changes over time across all three platforms. (B) Alamar alone showed significant changes in IL13, IL20, and IL4, which are all relatively low-abundance targets that had greater detectability for the Alamar assay, with the exception of IL4 for Luminex. (C) Alamar and Olink both showed significant changes in FLT3LG, but Luminex did not, even though detectability

715 was higher for Luminex than Alamar. (D) Alamar and Luminex both showed significant changes in IL5,
716 but Olink did not. Olink showed low detectability for IL5. (E) CCL3 was only significant for Olink, which
717 also had higher detectability than the other two platforms. (F) VEGFA was only significant for Luminex,
718 despite high detectability for all three platforms.
719