

1 Immunological and Antigenic Signatures Associated with Chronic Illnesses after 2 COVID-19 Vaccination

3 Bornali Bhattacharjee^{1,2#}, Peiwen Lu^{1#}, Valter Silva Monteiro^{1#}, Alexandra
4 Tabachnikova^{1#}, Kexin Wang^{2,11#}, William B. Hooper^{1,2#}, Victoria Bastos^{1#}, Kerrie
5 Greene^{1#}, Mitsuaki Sawano³, Christian Guirgis^{1,4}, Tiffany J. Tzeng^{1,2}, Frederick
6 Warner^{3,5}, Pavlina Baevova^{1,2}, Kathy Kamath⁶, Jack Reifert⁶, Danice Hertz⁷, Brianne
7 Dressen⁸, Laura Tabacof⁹, Jamie Wood⁹, Lily Cooke⁹, Mackenzie Doerstling⁹, Shadan
8 Nolasco⁹, Amer Ahmed⁹, Amy Proal^{9,10}, David Putrino⁹, Laying Guan^{2,11*}, Harlan M.
9 Krumholz^{2,3,12*}, Akiko Iwasaki^{1,2,13*}

10 Affiliations:

- 11 1. Department of Immunobiology, Yale University School of Medicine, New Haven, CT,
12 USA
- 13 2. Center for Infection and Immunity, Yale School of Medicine, New Haven, CT, USA
- 14 3. Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of
15 Medicine, Yale New Haven Hospital Center for Outcomes Research and Evaluation,
16 Yale New Haven Hospital, CT, USA
- 17 4. Department of Molecular, Cellular, and Developmental Biology, Yale College, New
18 Haven, CT, USA.
- 19 5. Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA
- 20 6. SerImmune, Goleta, CA USA
- 21 7. Independent Researcher, CA, USA
- 22 8. Independent Researcher, UT, USA
- 23 9. Cohen Center for Recovery from Complex Chronic Illness, Department of
24 Rehabilitation and Human Performance, Icahn School of Medicine at Mount Sinai, New
25 York City, NY, USA
- 26 10. Polybio Research Foundation, Boston, MA, USA
- 27 11. Department of Biostatistics, Yale School of Public Health, New Haven, CT, USA
- 28 12. Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School
29 of Medicine, New Haven, CT, USA
- 30 13. Howard Hughes Medical Institute, Chevy Chase, MD, USA

31

32 #: Indicates equal contribution

33 *: Co-senior and corresponding authors

34 Correspondence can be addressed to:

35 leying.guan@yale.edu, harlan.krumholz@yale.edu, akiko.iwasaki@yale.edu

36

37 **Keywords**

38 Post-vaccination syndrome, machine learning, immune cell populations, circulating

39 Spike, Epstein-Barr Virus, autoantibodies, antigen persistence

40

41 **SUMMARY**

42 COVID-19 vaccines have prevented millions of COVID-19 deaths. Yet, a small fraction
43 of the population reports a chronic debilitating condition after COVID-19 vaccination,
44 often referred to as post- vaccination syndrome (PVS). To explore potential
45 pathobiological features associated with PVS, we conducted a decentralized, cross-
46 sectional study involving 42 PVS participants and 22 healthy controls enrolled in the
47 Yale LISTEN study. Compared with controls, PVS participants exhibited differences in
48 immune profiles, including reduced circulating memory and effector CD4 T cells (type 1
49 and type 2) and an increase in TNF α + CD8 T cells. PVS participants also had lower
50 anti-spike antibody titers, primarily due to fewer vaccine doses. Serological evidence of
51 recent Epstein-Barr virus (EBV) reactivation was observed more frequently in PVS
52 participants. Further, individuals with PVS exhibited elevated levels of circulating spike
53 protein compared to healthy controls. These findings reveal potential immune
54 differences in individuals with PVS that merit further investigation to better understand
55 this condition and inform future research into diagnostic and therapeutic approaches.

56

57

58 INTRODUCTION

59 The rapid development and deployment of COVID-19 vaccines have been pivotal in
60 mitigating the impact of the pandemic¹. These vaccines have significantly reduced
61 severe illness and mortality associated with SARS-CoV-2 infection². Additionally,
62 vaccinated individuals experience a lower incidence of post-acute sequelae of COVID-
63 19 (PASC) or long COVID, thus highlighting an additional potential benefit of receiving
64 the COVID-19 vaccines^{3,4}. However, COVID-19 vaccines are associated with rare
65 acute adverse events⁵ such as myocarditis and pericarditis⁶, thrombosis and
66 thrombocytopenia⁷, Guillain–Barre syndrome, transverse myelitis, and Bell’s Palsy^{8,9}.

67 In addition, some individuals have reported post-vaccination symptoms resembling long
68 COVID beginning shortly after vaccination. This condition, sometimes referred to as
69 post-vaccination syndrome (PVS) or post-acute COVID-19 vaccination syndrome
70 (PACVS)^{10,11}, is characterized by symptoms such as exercise intolerance, excessive
71 fatigue, numbness, brain fog, neuropathy, insomnia, palpitations, myalgia, tinnitus or
72 humming in ears, headache, burning sensations, and dizziness¹⁰. Unlike long COVID,
73 PVS is not officially recognized by health authorities, which has significantly limited
74 patient care and support.

75 The molecular mechanisms of PVS remain largely unknown. However, there is
76 considerable overlap in self-reported symptoms between long COVID and PVS, as well
77 as shared exposure to SARS-CoV-2 spike (S) protein in the context of inflammatory
78 responses during infection or vaccination^{10,12,13}. In susceptible individuals, vaccines may
79 contribute to long-term symptoms by multiple mechanisms. For example, vaccine
80 components, such as mRNA, lipid nanoparticles, and adenoviral vectors, trigger
81 activation of pattern recognition receptors^{14,15}. Thus, unregulated stimulation of innate
82 immunity could lead to chronic inflammation. Secondly, it has been shown that the S
83 protein expressed following BNT162b2 or mRNA-1273 vaccination circulates in the
84 plasma as early as one day after vaccination^{16,17}. Interaction with full-length S, its
85 subunits (S1, S2), and/or peptide fragments with host molecules may result in
86 prolonged symptoms in certain individuals¹⁶. Recently, a subset of non-classical
87 monocytes has been shown to harbor S protein in patients with PVS¹⁸. Further,
88 biodistribution studies on mRNA–LNP platforms in animal models indicate its ability to
89 cross the blood-brain barrier, and the local S expression could result in neurocognitive
90 symptoms^{19,20}. Third, vaccine-induced immune responses may be triggering the
91 stimulation of autoreactive lymphocytes²¹.

92 To investigate immunological features in people suffering from persistent symptoms
93 after COVID-19 vaccination, a cross-sectional case-control study was undertaken to
94 identify the immunological correlates of PVS. A total of 42 participants with PVS who
95 had no pre-existing comorbidities and 22 contemporaneous healthy controls who did not
96 report PVS after receiving COVID-19 vaccines were included. An important factor to
97 evaluate was the possibility that PVS might result from an undiagnosed, asymptomatic
98 SARS-CoV-2 infection coinciding with the vaccination period, instead of being directly
99 caused by the vaccine administration. In addition, infection with SARS-CoV-2
100 significantly impacts immune signatures²². Our objectives were twofold: (1) to conduct a

101 two-group case-control analysis of the immunophenotypic profiles of individuals with
102 PVS in comparison with asymptomatic vaccine recipients, and (2) to compare the
103 immunophenotypic profiles of those with PVS with or without a history of SARS-CoV-2
104 infection. To achieve these, we profiled circulating immune cell populations, antibody
105 responses, and circulating immune modulator levels in addition to assessing the
106 demographic and general health characteristics of the participants.

107 **RESULTS**

108 **Cohort Description**

109 All the blood samples were collected between December 2022 and November 2023
110 from the Listen to Immune, Symptom and Treatment Experiences Now (LISTEN) study
111 ²³. The PVS cohort consisted of a total of 42 participants, including 29 females and 13
112 males with no preexisting comorbidities, whereas the control cohort consisted of 22
113 participants with 11 from each biological sex (Figures S1A and 1A). Upon recruitment of
114 44 PVS participants, two had to be excluded from the analyses due to evidence of
115 pharmacological immunosuppression. Information on index vaccine types was reported
116 by 39 out of 42 PVS participants included in the analyses, and they were Comirnaty
117 (Pfizer) (n=14), Spikevax (Moderna) (n=21), and Jcovden (J&J) (n=4). The most
118 frequent symptoms reported by participants were excessive fatigue (85%), tingling and
119 numbness (80%), exercise intolerance (80%), brain fog (77.5%), difficulty concentrating
120 or focusing (72.5%), trouble falling or staying asleep (70%), neuropathy (70%), muscle
121 aches (70%), anxiety (65%), tinnitus (60%) and burning sensations (57.5%). Further,
122 pairwise Euclidean distances were calculated in a sex-segregated manner based on the
123 presence of symptoms at recruitment and two distinct clusters of symptoms were
124 identified in both (Table S1, Figure S1B).

125 Between the case and control cohorts, a total of 15 (35.7%) and 10 (45.5%) reported
126 having a history of one or more previous SARS-CoV-2 infections, respectively (Table
127 S2). However, upwards of 40% of SARS-CoV-2 infections are asymptomatic²⁴. To
128 further investigate prior history of SARS-CoV-2 infections, plasma specimens were
129 analyzed using the EUA-cleared Elecsys® anti-SARS-CoV-2 immunoassays, which
130 measure the presence of high affinity IgM, IgA, and IgG anti-N antibodies. A cut-off
131 index ≥ 1 was defined as reactive based on previous literature²⁵. In the non-reactive
132 group, the antibody indices varied between 0.09 and 0.17, whereas in the reactive
133 group, they ranged between 1.37 and 94.4. Among participants with PVS, 26 (61.9%)
134 were found to be reactive compared with 10 (45.5%) among controls. One participant
135 from each cohort with self-reported history of infection had non-reactive test results.
136 Based on both self-reports and serological analyses the two cohorts were further
137 classified into four subgroups, PVS with no history of infection (PVS-I, n= 15), PVS with
138 a history of infection (PVS+I, n= 27), controls with no history of infection (Control-I, n=
139 11) and controls with a history of infection (Control+I, n= 11) (Figure 1A, Table S2).
140 Even though all PVS participants developed chronic symptoms following vaccination
141 and not infection, it was important to consider the impact of a subsequent SARS-CoV-2
142 infection on immune phenotypes analyzed in this study.

143 Among the demographic variables, there were no significant differences in the number
144 of males and females between cases and controls (Fisher's Exact test, $p = 0.58$) or
145 among the four groups (Kruskal-Wallis test, $p = 0.25$; Figure 1B). Similarly, no significant
146 age differences were observed between cases and controls, median age (42.5 years,
147 PVS; 38 years, controls, Mann-Whitney U test $p = 0.27$) and among the four groups
148 ($p = 0.17$; Figure 1C).

149 The self-reported General Health Visual Analogue Scale (GHVAS) scores on the day of
150 biospecimen collection differed significantly among the four groups (Kruskal-Wallis test,
151 $p < 0.01$). The controls in both subgroups had significantly higher median scores
152 compared with the PVS subgroups (64, PVS-I; 60, PVS+I; 95, Control-I; 90, Control+I;
153 Figure 1D). The PROMIS-29 physical function, fatigue, pain interference, depression,
154 anxiety, sleep disturbance, and pain interference scores were compared independently
155 among the four groups to gauge the physical and mental health status of the
156 participants. The physical function scores were significantly higher among the controls
157 than the cases irrespective of infection status, median scores (13, PVS-I; 14.5, PVS+I;
158 20, Control-I; 20, Control+I). The anxiety (9.5, PVS-I; 10, PVS+I; 5, Control-I; 4,
159 Control+I), depression (8.5, PVS-I; 8, PVS+I; 4, Control-I; 4, Control+I), fatigue (16,
160 PVS-I; 15, PVS+I; 6, Control-I; 7, Control+I) and pain scores (9.5, PVS-I; 12, PVS+I; 4,
161 Control-I; 4, Control+I) were significantly lower in controls compared with participants
162 with PVS irrespective of infection status. Further, significantly higher sleep disturbance
163 scores were observed only among infection-negative cases compared to control-I
164 participants (13, PVS-I; 10, Control-I) (Figure 1E).

165 Most individuals in each cohort completed the primary series of vaccines based on
166 WHO recommendations (83.3%, PVS; 100%, Controls; Fisher Exact test, $p = 0.09$).
167 Participants with PVS received significantly fewer COVID-19 vaccine doses compared
168 with controls, median vaccine numbers (2, PVS; 4, controls; Fisher Exact test,
169 $p < 0.01$). On similar lines, the median number of days post the latest vaccination was
170 significantly higher among cases with a median of 585 days (± 190) compared with 199
171 days (± 217) among controls (Mann-Whitney U test, $p < 0.01$). In 85% of the cases,
172 participants identified the index vaccine dose as being part of the primary series [dose
173 1(45%) and dose 2(40%); Figure 1F]. The median number of days for the development
174 of any symptom was 4 [Interquartile range (IQR): 23 days], while for severe symptoms,
175 it was 10 (IQR: 44 days) post-vaccination. A high proportion of participants with PVS
176 developed any symptoms (70%) or severe symptoms (52.2%) within 10 days of
177 vaccination (Figure 1G).

178 Differences in circulating immune cell populations

179 To determine immune signatures of PVS, peripheral blood mononuclear cells (PBMC)
180 were analyzed using flow cytometry. Among the cell populations of myeloid lineage,
181 proportions of non-classical monocytes ($CD14^{low}CD16^{high}$; Mann-Whitney U test, $p =$
182 0.03) were significantly higher in the PVS cohort compared to the controls without
183 significant differences in the percentage of total monocytes despite greater median
184 values in PVS (Figure 2A). The median percentage of conventional type 2 dendritic cells
185 (cDC2; $CD304^+HLA-DR^+/CD1c^+$) was significantly lower among the participants with

186 PVS compared to the controls ($p= 0.02$) while no differences were observed in the
187 proportions of conventional type 1 dendritic cells (cDC1; $(CD304^-/HLA-DR^+/CD141^+)$
188 (Figure 2B). Pairwise comparisons were also executed to understand the differences
189 among the PVS subgroups with or without a history of infection. Among the low-density
190 granulocytes, no differences were observed between cases and controls in the
191 proportions of eosinophils ($CD66b^+CD56^-CD16^-$) or between the cases and controls in
192 the infected or uninfected subgroups but the proportion of neutrophils ($CD66b^+/CD56^-$
193 $/CD16^+$) was significantly higher ($p= 0.02$) in infection positive PVS subgroup (PVS+I)
194 compared to the convalescent controls (control+I) (Figure S2A). Significantly lower and
195 higher proportions of classical and non-classical monocytes, respectively, were
196 observed among the PVS+I compared to the control+I subgroup ($p(\text{cMonocytes})=$
197 <0.01 ; $p(\text{ncMonocytes})= 0.03$) with no differences between the cases and controls
198 without prior history of SARS-CoV-2 infection (Figure S2B). Next, significantly higher
199 proportions of both cDC1 and cDC2 cells were observed in the control+I subgroup
200 compared to the PVS+I subgroup ($p= 0.03$ and $p= <0.01$), respectively (Figure S2C).

201 Among the B cell populations, relative proportions of unswitched memory B cells (US
202 memory B cells; $CD19^+/CD27^+/IgD^+$) were significantly higher ($p= 0.02$) while the
203 proportion of double negative B cells (DN B; $IgD^-/CD27^-/CD24^-/CD38^-$) was observed to
204 be lower ($p= 0.01$) in the PVS cohort compared with controls (Figure 2C).

205 Significant differences in subsets of circulating immune cell populations were observed
206 across T cell lineages. Upon assessment of the T cell populations, notably higher
207 proportions of effector memory CD4 T cell subsets ($CD4^+Tem$; $CD45RA^-/CD127^+/CCR7^-$
208 ; $p= 0.01$) and resting natural CD4⁺ Treg; $CD45RA^+/CD25^+/CD127^-/HLA-DR^-$; $p= 0.05$)
209 were observed among the controls (Figure 2D). However, the PVS cohort had
210 significantly higher proportions of exhausted CD8 T cell ($CD8^+Tex$; $PD-1^+/TIM3^+$; $p=$
211 0.02) (Figure 2F) with no observed differences in the CD4⁺ central memory ($CD4^+cm$;
212 $CD45RA^-/CD127^+/CCR7^-$) and exhausted ($CD4^+Tex$; $PD-1^+/TIM3^+$) CD4 T cell
213 populations (Figure 2D). Upon *in-vitro* stimulation, the expression of CXCR3 on the cell
214 surface (Mann-Whitney U test, $p= <0.01$), intracellular IL-4 levels ($p= 0.04$) and IL-4, IL-
215 6 in combination were found to be significantly lower in the CD4 T of PVS cohort ($p=$
216 <0.01), with no differences were observed in $IFN\gamma$ & $TNF\alpha$ levels (Figure 2E).
217 Significant increases in intracellular $TNF\alpha$ levels ($p= <0.01$) with non-significant
218 increases in $IFN\gamma$ in the stimulated CD8 T cells were observed in PVS cohort (Figure
219 2F). Only a total of 32.23% of the variability in intracellular $TNF\alpha$ could be explained by
220 $IFN\gamma$ levels in the CD8⁺ T cell populations ($R^2= 0.32$; Figure 2G).

221 In the subgroup analyses, no differences in proportions of DN B cell subpopulations
222 were observed (Figure S2D). Proportions of CD4⁺ CD45RA⁺ effector memory T cells
223 ($CD4^+T_{EMRA}$; $CD45RA^+/CD127^-/CCR7^-$; $p= 0.02$) and rTregs ($p= 0.03$) were both
224 observed to be significantly lower in PVS+I compared to the control+I subgroup (Figure
225 S2E). Proportions of CXCR3 expressing stimulated CD4 T cells was much lower in
226 PVS+I cases ($p= <0.01$) and the proportions of both IL-4⁺ ($p= 0.04$), and IL-4⁺/IL-6⁺ ($p=<$
227 0.01) double-positive cells were also lower compared to the controls (Figure S2E).
228 Higher proportion of CD8⁺ Tcm cells was retained in the PVS+I subgroup compared to

229 the control+I group (Figure S2F; $p= 0.02$). No differences were observed in immune cell
230 populations between the infection-negative cases (PVS-I) and controls (Controls-I).

231 **Lower levels of spike-specific antibody responses in PVS**

232 Given the differences in the number of vaccine doses received between participants in
233 the PVS cohort and the control group, we compared spike-specific immunoglobulin G
234 (IgG) levels in relation to the number of vaccine doses administered. Correlation
235 analyses revealed a significant positive correlation between the number of vaccine
236 doses and plasma anti-S IgG levels (Spearman's Rank Correlation Coefficient, $\rho = 0.85$,
237 $p= <0.01$) as well as anti-RBD IgG levels ($\rho = 0.83$, $p= <0.01$) in the PVS-I subgroup. In
238 the PVS+I subgroup, only anti-S IgG levels showed a significant correlation ($\rho = 0.55$, p
239 $= 0.01$) with the number of doses (Figure 3A). Next, correlation analyses were
240 performed to assess the relationships between plasma anti-S, anti-RBD, and anti-N IgG
241 levels with the number of days post last vaccination among the four groups. No
242 significant changes in anti-S and anti-RBD antibody levels were observed with
243 increasing days since vaccination in the control group, regardless of infection history,
244 and in the PVS+I subgroup (Fig 3B). In contrast, significant negative correlations were
245 found in the PVS-I subgroup between the number of days post-vaccination and both
246 anti-S ($\rho = -0.87$, $p= <0.01$) and anti-RBD ($\rho = -0.83$, $p= <0.01$) IgG levels, indicating a
247 decline in these antibodies over time (Fig. 3B). Additionally, as expected, no
248 correlations were observed between anti-N IgG levels and days post vaccination across
249 the infection-positive subgroups.

250 The next step was to evaluate if the most recent exposure to SARS-CoV-2 or
251 vaccination correlated with the observed differences in waning patterns. No significant
252 changes in anti-S, anti-RBD and anti-N antibody levels were observed with an increase
253 in the number of days from self-reported viral infection dates among the Control+I and
254 PVS+I subgroups (Figure 3C). In addition, the plasma titers of anti-S IgG were
255 significantly lower among the PVS-I cases compared to the Control-I subgroup ($p=$
256 <0.01) (Figure 3D). However, no differences were observed in the anti-RBD IgG levels
257 across the four subgroups (Figure 3D). As expected, the uninfected PVS-I and the
258 Control-I subgroups had much lower anti-N IgG levels as detected by in-house ELISAs
259 (Figure 3D). To further account for variations in vaccine doses and infection, we
260 developed linear models. Those models indicated that both prior SARS-CoV-2 infection
261 and the number of vaccine doses were significantly associated with higher levels of anti-
262 RBD and anti-S IgG (Figure 3E, Table S3).

263 **Serological evidence of recent EBV reactivation in PVS**

264 Many human pathogens are ubiquitous, opportunistic, and capable of establishing
265 lifelong infections with alternate latency and reactivation cycles²⁶. These cycles can be
266 triggered by physiological perturbations and can contribute to systemic inflammation²⁷.
267 Therefore, we used serum epitope repertoire analysis (SERA) to evaluate seropositivity
268 against a range of pathogens, including five bacterial, seven parasitic, 14 viral and one
269 fungal species. On performing two group analyses, no significant differences were
270 observed for all pathogens, indicating similar levels of prior exposure. (Fig. 4A).

271 Moreover, the seropositivity for each pathogen did not significantly differ from
272 seropositivity in 3448 healthy controls collected before the COVID-19 pandemic (Figure
273 4A). Given the high seropositivity rates for herpesviruses, we further analyzed the
274 seropositivity patterns in combination, for cytomegalovirus (CMV), Epstein-Barr Virus
275 (EBV), Herpes Simplex Virus Type 1 (HSV-1) and Herpes Simplex Virus Type 2 (Figure
276 4B). Significant differences were observed between cases and controls (Mann-Whitney
277 U test, $p = 0.01$; Figure 4C), where the participants with PVS had higher prevalence of
278 EBV and HSV coinfection, and lower prevalence of EBV and CMV coinfection. There
279 are reports of similarities in symptom phenotypes between PVS and long COVID, as
280 well as evidence of EBV reactivation in long COVID cases, including elevated
281 antibodies against EBV surface protein gp42^{22, 28}. Therefore, we further investigated the
282 prevalence of antibodies against EBV gp42 and identified significantly elevated
283 antibodies in the plasma of PVS participants compared with controls (Kruskal-Wallis
284 test, $p = <0.01$, Figure 4D, E). As an orthogonal validation, we tested the distribution of
285 linear peptide reactivities across the EBV proteome. Greater reactivities to two peptides
286 corresponding to two envelope glycoproteins necessary for B cell infection, gp42 and
287 gp350 were observed. For the gp42 protein, the antibody reactivity to peptide
288 ([VI]XLPHW) was significantly higher among the PVS participants irrespective of the
289 SARS-CoV-2 infection status (Mann Whitney U test, $p = <0.01$; Figure 4F) and across
290 the four subgroups (Kruskal-Wallis test, $p = 0.02$; Figure 4G). Greater reactivities were
291 also observed for the gp350 peptide (KXR[X]RQ]W[X]F) among the PVS participants
292 compared to controls (Kruskal-Wallis test, $p = 0.04$; fig 4J) and across the four
293 subgroups (Kruskal-Wallis test, $p = 0.03$; Figure S3C). Further, anti-gp42 ([VI]XLPHW)
294 reactivity by SERA significantly correlated with anti-gp42 ELISA measurements thus
295 validating the finding ($R = 0.37$, $p = <0.01$; Figure 4I). We also mapped this motif onto
296 available structures of gp42 complexed with EBV gH/gL (PDB: 5T1D), demonstrating its
297 location close to the transmembrane (TM) domain of gp42 and surface-exposed (Figure
298 4H). Study participants with greater antibody reactivity to gp42 as assessed by ELISA
299 also exhibited higher percentages of TNF α -producing CD8⁺ T-cells ($R = 0.47$, $p = <0.01$,
300 Figure 4K). This correlation was not observed for IL-4, IL-6 double-positive CD4 T cells
301 (Figure S3D) as was previously reported for long COVID²².

302 **Participants with PVS have a distinct set of autoantibodies**

303 To evaluate differences in immunoglobulin isotypes and IgG subtypes in the plasma,
304 Luminex assays were performed. No significant differences were observed between the
305 PVS cohort and the controls (Figure S4A). Next, to determine the presence of
306 autoantibodies in PVS, we screened for reactivities across a range of 120 known
307 autoantigens using microarrays for three different immunoglobulin isotypes, IgM, IgG,
308 and IgA. We observed significant increases in IgM reactivities against 65 antigens, IgG
309 reactivity against 1 antigen and IgA reactivities against 39 antigens in PVS compared to
310 controls after multiple testing corrections (Table S4). Among these antigens, two
311 showed log₂fold change of greater than 2: anti-nucleosome IgM (Mann-Whitney U test,
312 $p = <0.01$) and anti-AQP4 IgA ($p = <0.01$) (Figure S4B). Conversely, control participants
313 exhibited higher reactivities against a total of 21 antigens, 18 of which were of the IgG
314 isotype and five were of IgA isotype with two common antigens between the two
315 isotypes (Table S3). Among these autoantibodies, anti-histone H1 IgG differed by

316 greater than log₂fold change ($p = <0.01$, Figure S4B). Infection-positive subgroups had a
317 higher number of reactivity differences between cases and controls (Figure S4C).
318 Among the PVS-I participants, anti-calprotectin/S100 IgM, anti-genomic DNA IgA and
319 anti-ssDNA IgA reactivities were significantly higher while anti-histone H3 IgG, anti-MBP
320 IgA, and anti-PR3 IgA reactivities were higher among the controls-I (Figure S4C, Table
321 S5).

322 **Circulating hormones and immune modulators in PVS**

323 Two group analyses of circulating hormones and immune modulators revealed
324 significantly lower levels of fetuin A26 and neurotensin (Mann-Whitney U tests, $p = 0.01$
325 and $p = 0.03$; Figure S5A) in participants with PVS with fold differences of 1.3 and 1.9
326 respectively. Additionally, four group analysis was performed to evaluate the impact of
327 infection on PVS. Given the smaller number of samples in the four group analyses,
328 each panel of analytes was independently evaluated. Significantly lower levels of
329 circulating fetuin A36, and neurotensin were also observed among participants with
330 PVS with a history of SARS-CoV-2 infection compared to convalescent participants ($p =$
331 0.01 for both analytes; Figures S5B-C). No differences were observed for other factors
332 across the subgroups except for β endorphin which was significantly lower in PVS+I
333 compared to the control+I group ($p = 0.01$; Table S7) without any significant differences
334 in the two group analyses. No differences were observed in the uninfected subgroups.

335 **Increase in circulating SARS-CoV-2 Spike protein in participants with PVS**

336 It has been reported that the BNT162b2 or mRNA-1273 derived S proteins circulate in
337 the plasma of those vaccinated as early as one day after the vaccine and interactions of
338 the circulating protein¹⁶. Hence, we next sought to investigate whether the S1 subunit of
339 the SARS-CoV-2 S protein could be detected in the plasma. For this, we used an anti-
340 S1 Successive Proximity Extension Amplification Reaction (SPEAR) immunoassay.
341 This method can detect S1 levels as low as 5.64 fM. We conducted a one-sided
342 Kolmogorov–Smirnov test with 1000 permutations to see if the participants with PVS
343 had higher circulating S1. The results indicated that participants with PVS had
344 significantly higher circulating S1 levels compared with the control group ($p = 0.01$).
345 However, circulating S1 was found in only a subset of participants with PVS at varying
346 concentrations while the control group mostly exhibited a bimodal distribution of zero
347 and non-zero values (Fig. 5A, Table S2). Detectable S1 was found in participants'
348 plasma ranging from 26 to 709 days from the most recent known exposure (Figure 5B).
349 To fully account for the width of this dataset, we included all non-detectable values in
350 the analysis and applied a generalized regression model accounting for zero-inflation.
351 We found that both PVS-I and PVS+I groups displayed significantly elevated S1 levels
352 than the Control-I group ($p = <0.01$ and $p = 0.02$, respectively) (Figure 5C).

353 Given the similarities between PVS and long COVID symptoms, one hypothesis in the
354 literature is that shared exposure to the S protein may play a role and several groups
355 have independently reported the presence of circulating S1 & full-length S in long
356 COVID using various detection methods^{16,29}. To further investigate this circulating S1
357 positivity percentages and levels in the LISTEN cohort subgroups were compared with

358 an external cohort of healthy, convalescent controls and LC participants (MY-LC
359 cohort). This external cohort, collected from Mount Sinai clinics, included 134
360 healthy/convalescent controls and 134 long COVID participants, which were all assayed
361 together with the LISTEN cohort biospecimens.

362 Among the MY-LC healthy [HC(n= 62); no reported SARS-CoV-2 infection],
363 convalescent controls [CVC(n=72); with reported SARS-CoV-2 infection], and long
364 COVID participants, detectable S1 was observed in 30.6% of control participants
365 (41/134; HC= 12.9%; CVC= 22.2%) and 33.6% (45/134) of individuals with long COVID,
366 with mean S1- $\ln(x_{fM} + 1)$ values of 3.72 and 3.85 respectively among those above the
367 LLoD. These figures were comparable to the percentages observed among the LISTEN
368 controls (31.8%) and PVS (35.7%) groups. S1 levels were moderately elevated in the
369 MY-LC control group compared to the PVS control group ($p=0.06$), potentially reflecting
370 differences in exposure timing or SARS-CoV-2 variant of concern (VOC). Despite this,
371 the PVS group demonstrated significantly higher S1 levels compared to both control
372 cohorts (LISTEN-control: $p<0.01$; MY-LC control: $p=0.03$; mean S1- $\ln(x_{fM} + 1) = 6.24$)
373 (Fig. 5D, Table S8).

374
375 To further validate the findings and to investigate whether the presence of S1 reflects
376 the presence of full-length S protein among the LISTEN participants, we next conducted
377 a full-length S SPEAR assay. The calculated values for the LLoD and the LLoQ were
378 1.81 and 8.24 fM, respectively. The full-length S SPEAR assays showed a significant
379 correlation between S and S1 across all samples (Fig. 5E), as well as for values above
380 the LLoD and LLoQ (Fig. 5F, Table S2). Thus, based on SPEAR assays, the individuals
381 with PVS exhibited elevated levels of circulating full-length S compared to healthy
382 controls.

383 **Immune signatures in PVS subgroups based on the presence of circulating S1** 384 **protein**

385 To gain a clearer understanding of the variability of circulating S1 protein levels, we first
386 compiled a structured timeline that summarizes the self-reported infection dates,
387 vaccine numbers (including types and administration dates), and the number of days
388 between the latest known exposure and the collection of biospecimens. This timeline
389 was organized for both the PVS-I and PVS+I groups (Figure 5G). Notably, we observed
390 that the highest levels of detectable S1 in the PVS-I group were the furthest away from
391 the last known exposure and ranging between greater than 600-700 days (NI-1 & NI-5;
392 Figure 5G). This suggested that prolonged antigen persistence might be associated with
393 PVS in a subgroup of patients. Further, most of the PVS+I group participants
394 experienced breakthrough SARS-CoV-2 infections with the exception of two cases,
395 indicating that PVS symptoms started prior to infection (Figure. 5G).

396 Given the possible heterogeneity in immunological trajectories leading up to PVS and
397 the lack of adequate sample numbers in each PVS subgroup, we next took a more
398 descriptive approach to look for peripheral immune signatures stratified based on their
399 infection status and detectable S1 above SPEAR assay's lower limit of quantitation. In
400 order to begin with a valid method of selection despite the small sample sizes, non-

401 parametric Mann-Whitney tests were implemented without multiple testing corrections to
402 look for differences in distributions of 547 independent variables including GHVAS
403 scores, circulating modulator levels, anti-SARS-CoV-2 antibody titers and autoantibody
404 scores within both the PVS-I and PVS+I subgroups. Variables showing significant
405 differences were further filtered based on greater than 1.5 fold changes to identify the
406 distinct determinants associated with each of the four trajectories.

407 Among other factors, the infection-naïve PVS participants with quantifiable S1 had lower
408 GHVAS scores indicative of poorer general health (GHVAS) and lower anti-S IgG titers,
409 whereas higher circulating IL-7 and IL-21 levels were detected compared to other
410 groups (Figure. 5H). Elevated growth hormone levels alongside low TSH levels were
411 also observed among the PVS-I participants with S1 protein in circulation. By contrast,
412 among the infection-positive subgroups, participants with circulating S1 were observed
413 to have higher anti-N antibody titers based on the clinical COBAS assays and in-house
414 ELISA indicative of the contribution of infection history. Moreover, anti-nucleosome IgA
415 levels were higher among those in the PVS+I subgroup without detectable S1 (Figure.
416 5H).

417 **Machine Learning-based identification of peripheral immune signatures of PVS**

418 To establish a combined global immune signature for persistent symptoms following
419 COVID-19 vaccination, we built machine learning models to predict PVS outcomes. The
420 goal was to identify prominent features that could effectively distinguish PVS from the
421 controls in a parsimonious manner. Given that autoantibodies are also common in the
422 general population at low levels, we chose to exclude them from this analysis in the
423 absence of further validation³⁰. Additionally, we excluded any variables with greater than
424 20% missing values for either the PVS or control groups and SERA variables because
425 most of the dataset lacked a significant number of values above LLoD. A total of 193
426 variables were included.

427 Weighted Gene Co-Expression Network Analysis (WCGNA)³¹ was applied to this set to
428 find groups of highly correlated variables (Figure S6B). The final feature set was then
429 created by taking all variables that did not form a tight cluster (141, Module 1) and the
430 eigengene, or first principal component, of every other set of variables (Modules 2-6).
431 This gave us a total of 146 features. Next, we performed classification utilizing Least
432 Absolute Shrinkage and Selection Operator (LASSO) with nested cross-validation. We
433 achieved an overall model accuracy of 78.1% on the validation folds (per-fold range
434 62.5% - 100%; Figure S6A), and an AUC of 0.80 (per-fold range 0.67-1.00, 95% CI =
435 0.67-0.92; Figures 6A and S6A). A permutation test further concluded that this
436 performance was significantly above random ($p = 0.02$; Figure S6A). Segregating test
437 data per-fold by infection status yielded accuracy on the infected population of 86.5%
438 and the uninfected population of 73.3%. (Figure S6A) Per-class accuracy showed some
439 divergence, with PVS accuracy of 85.7% vs. control accuracy of 63.6%. This was due to
440 a subset of the controls clustering primarily with PVS samples, making control
441 classification more difficult (Figure 6B).

442 The LASSO model selected 21 features using all data, consisting of CD4 T cell
443 populations, immune modulators, neuropeptides, and antibodies (Figure 6C). Among
444 the features selected, there were several negatively associated with PVS. These
445 included circulating factors sIL-1R1, fetuin A36, granzyme A and B, FLT-3L and
446 HMGB1, and subsets of circulating CD4 T cell populations (CXCR3⁺ CD4 T cells CD4⁺
447 T_{EMRA} cells, and IL-4⁺/IL-6⁺ CD4 T cells). Multiple hormones and neuropeptides
448 synthesized by the hypothalamus, pituitary glands, and the peripheral nerves and
449 involved in nociception and stress responses such as oxytocin, neurotensin, β
450 endorphin, melanocyte-stimulating hormones (MSH), and substance P were also
451 negatively associated with PVS and formed a single module (Module 6) (Figure 6D).
452 The features that were positively associated with PVS were anti-EBV gp42 IgG titers,
453 MMP1 levels, and TNF α ⁺ CD8 T cells. We observed that no single variable or small
454 subset of variables had a particularly strong differentiating power.

455 DISCUSSION

456 In this study, we examined symptoms and circulating immune factors and cell types
457 associated with chronic illness following COVID-19 vaccination. Post-acute conditions
458 following COVID-19 vaccination have been reported for multiple vaccine platforms
459 including mRNA and adenoviral-vectored vaccines^{6,7,8,9}. We observed that the general
460 health status of the PVS participants was far below the general US population average³²
461 based on the GHVAS scores. The patient-reported outcome scores from the
462 PROMIS29 domains were also indicative of lower quality of life. A large fraction of
463 individuals reported the onset of symptoms to be as early as within one day of COVID-
464 19 vaccination. Compared with controls, participants with PVS had reduced CD4⁺ T cell
465 subsets in circulation (both Th1 and Th2) and an increased percentage of TNF α ⁺ CD8 T
466 cells. Among cell populations of myeloid origin, cDC2 cells were reduced, and non-
467 classical monocytes were elevated among PVS participants. Lower S-specific IgG
468 levels were observed in PVS mainly due to the limited vaccine doses received.
469 Additionally, serological evidence for recent EBV reactivation was also observed. Using
470 machine learning approaches, we further identified a set of 21 core predictive features
471 of PVS status within the LISTEN PVS cohort with potential for further validation and
472 biomarker identification. Most notably, we found elevated levels of spike (S1 and full-
473 length S) in circulation up to 709 days after vaccination among a subset with PVS, even
474 in those with no evidence of detectable SARS-CoV-2 infection.

475 To date, only a few studies have investigated the immunological mechanisms
476 associated with PVS^{11,12,33}, and no consensus definition of this syndrome exists^{10,34}.
477 Previous studies on PVS have found the presence of elevated levels of inflammatory
478 cytokines such as CCL5, IL-6, and IL-8; IgG subclass imbalances, high angiotensin II
479 type 1 receptor antibodies (AT1R), and the presence of spike S1 in non-classical
480 monocytes, among others^{11,12,33}. In the LISTEN PVS cohort, we did not find evidence of
481 elevation in inflammatory cytokines or IgG subclass imbalances. This difference may be
482 due to the heterogeneity of the cohorts studied, vaccine types or the time from
483 vaccination.

484 The demographics at risk of developing PVS and symptom manifestations are similar to
485 those of long COVID^{10, 35,36,37}. Whether this reflects overlapping underlying
486 mechanisms such as persistent S protein remains to be determined. Circulating S1
487 antigen has been detected in mRNA-1273 vaccine recipients without a prior history of
488 viral infection within an average of five days after the first injection and becomes
489 undetectable by day 14¹⁶. By contrast, in our study, significantly elevated levels of
490 circulating S1 and S were observed in a subset of PVS participants both in the infection-
491 naive and infection-positive groups up to 709 days post-exposure. This is in line with the
492 findings of S1 persistence in monocytes in people with PVS¹². Circulating full-length S
493 has also been detected in cases of post-vaccination myocarditis³⁸. Given the striking
494 similarities between long COVID and PVS symptoms, there has been speculation
495 regarding the potential causal role of the persistent presence of spike protein³⁹ driving
496 the chronic symptoms. Additionally, a recent study has shown spike protein binding to
497 fibrin resulting in inflammation *ex vivo* and neuropathy in animal experiments⁴⁰. S1
498 subunit is sufficient to cause formation of trypsin-resistant fibrin clots when added to
499 plasma from healthy individuals⁴¹. The persistent presence of S1 and the full-length
500 spike protein across multiple long COVID cohorts lends further support to this
501 hypothesis^{42,43,44,45}. Additionally, our results using the S1 SPEAR assays indicate higher
502 percentages of individuals with S1 antigen persistence among both MY-LC controls and
503 the long COVID group compared to other studies despite reporting mild acute phase
504 symptoms reported by these participants^{44,45,43}. This may be attributed to variations in
505 assay sensitivity or variations in vaccine doses and re-infection rates across the
506 cohorts. Despite higher antigen persistence rates, the PVS participants with detectable
507 S1 had higher mean circulating S1 levels compared to the LC participants. In our PVS-I
508 group, anti-S antibody levels were lower in those with circulating S1. Why persistent
509 spike antigen fails to elicit an antibody response, and what the source of persistent
510 spike in circulation is, requires further investigation.

511 Immunophenotyping of circulating PBMCs from participants with PVS revealed lower
512 levels of circulating CD4⁺ Tem, CXCR3 expressing CD4, as well as IL-4⁺/IL-6⁺ double
513 positive CD4 T cell populations and higher TNF α secreting CD8 cell populations. This is
514 in contrast to our observations of higher levels of IL-4⁺/IL-6⁺ CD4 T cell populations in
515 the long COVID cohort²². Elevated levels of anti-S IgG have been observed in long
516 COVID patients, possibly reflecting persistent S protein^{22,42}. By contrast, within the PVS-
517 I subgroup, the lower levels of anti-S antibodies were associated with a reduced number
518 of vaccinations. Moreover, PVS participants in this study did not exhibit decreased
519 circulating cortisol levels or increased fetuin A36 levels, as reported for long COVID^{22,46}.

520 While our panel of autoantigens did not include any G protein-coupled receptors
521 included in the Semmler et al study¹¹, among those identified to be elevated in PVS in
522 this study, namely, anti-nucleosome IgM and anti-AQP4 IgA require further
523 investigation. Higher monomeric IgM has been reported in autoimmune disease
524 patients and circulating nucleosomes have also been shown to trigger cGAS (cGMP-
525 AMP synthase) immune responses^{47,48}. Along similar lines, anti-AQP4 IgG is most
526 commonly associated with neuromyelitis optica spectrum disorder (NMOSD), however
527 there are not any reports on IgA isotype⁴⁹. In addition, similar to what has been reported

528 in long COVID, elevated antibody responses against EBV lytic antigen were detected
529 among seropositive participants with PVS, suggesting recent reactivation,^{22,28}.

530 This study has several limitations. Our small sample size could have affected the
531 robustness of the machine learning approaches and prediction of specific immune
532 features in PVS. Due to the limited sample size, we might have failed to capture small
533 but potentially important immune features associated with PVS. Analysis of autoimmune
534 antibody reactivity was restricted to antigens reported in other autoimmune diseases,
535 limiting the discovery of a broader range of autoantibodies. While we used two
536 independent approaches to ascertain previous infection with SARS-CoV-2, negative
537 results cannot definitively preclude prior infection that occurred in the distant past. Other
538 limitations include the lack of analysis of the host genetics that might account for PVS
539 susceptibility, or any other conditions, such as non-prescription drugs or asymptomatic
540 infection with other pathogens that were not tested in our analysis might have
541 predisposed an individual to develop chronic illness following COVID-19 vaccination.
542 While we observed elevated levels of S1 among those with PVS compared to LC,
543 additional studies with matched patient demographic profiles are necessary to
544 determine whether this represents genuine differences or is simply a result of random
545 variation. Finally, we do not know whether our findings extend beyond COVID-19
546 vaccination since we did not include PVS following other vaccines.

547 In summary, by revealing distinct immunological features of PVS, this study helped
548 generate hypotheses regarding the underlying pathobiology of this condition.
549 Understanding such mechanisms will help improve the overall safety profile of COVID-
550 19 vaccines and support public health strategies that maximize vaccine efficacy while
551 minimizing adverse effects. However, this study is early-stage and requires replication
552 and validation. We emphasize the critical task of discerning between meaningful results
553 and random fluctuations in the data. Future work is essential to elucidate these
554 relationships. As the global community continues to navigate the challenges of COVID-
555 19 and long COVID, a deeper understanding of vaccine-related immune responses will
556 be essential in refining vaccination practices and ensuring their long-term success.

557 **RESOURCE AVAILABILITY**

558 The flow data repository ID is FR-FCM-Z8FZ for all the raw .fcs files generated for flow
559 cytometry analyses at the Flow Repository platform⁵⁰. Custom codes used for
560 computational analyses will be made available by authors upon reasonable request.

561 **ACKNOWLEDGEMENTS**

562 We thank all the participants from the LISTEN study who contributed their time, effort
563 and precious biospecimens for this study. We thank the Yale Center for Clinical
564 Investigation (YCCI) coordinators Geisa Wilkins, Iris Dias and Yvette Strong for their
565 support. Victoria Bastos' fellowship was supported by CAPES Foundation, Ministry of
566 Education, Brazil. This study was funded in part by the Howard Hughes Medical
567 Institute Collaborative COVID-19 Initiative.

568 **AUTHOR CONTRIBUTIONS**

569 Experimental conceptualization, methodology and data visualization were performed by
570 B.B, P.L.,V.B., V.S.M., A.T., K.W. W.B.H, K.G., K.K. and J.R. Formal analysis was
571 conducted by B.B, P.L.,V.B., V.S.M., A.T., K.W. W.B.H and K.G. Resources were
572 provided by H.M.K and A.I. Review of survey responses and electronic health records
573 were performed by B.B. Sample collection, processing and biospecimen validation were
574 performed by B.B., T.J.T, P.B. and C.G., The original draft was written by B.B. and A.I.
575 Review and editing were performed by B.B., P.L.,V.B., V.S.M., A.T., K.W. W.B.H, K.G.,
576 K.K., J.R., D.H., B.D., L.G, H.M.K, and A.I. Data curation was performed by B.B.,
577 W.B.H, F.W. L.G. M.S., and H.M.K. and A.I. supervised the study. Funding was
578 acquired by A.I.

579 **DECLARATION OF INTERESTS**

580 In the past three years, H.M.K. received expenses and/or personal fees from United
581 Health, Element Science, Eyedentifeye and F-Prime; he is a co-founder of Refactor
582 Health, HugoHealth and MedRxiv; and is associated with contracts, through Yale New
583 Haven Hospital, from the Centers for Medicare & Medicaid Services and through Yale
584 University from the Food and Drug Administration, Johnson & Johnson, Google and
585 Pfizer. A.I. co-founded and consults for RIGImmune, Xanadu Bio and PanV and is a
586 member of the Board of Directors of Roche Holding and Genentech.

587 **SUPPLEMENTAL INFORMATION**

588 Document S1. Figures S1-S6

589 Excel file S2 containing additional data.

590 Table S1: Prevalence of PVS symptoms at the time of recruitment

591 Table S2: Demographic, clinical and immunological data on all participants included in
592 this study

593 Table S3: Generalized linear modeling analyses results

594 Table S4: Two group autoantibody analysis results

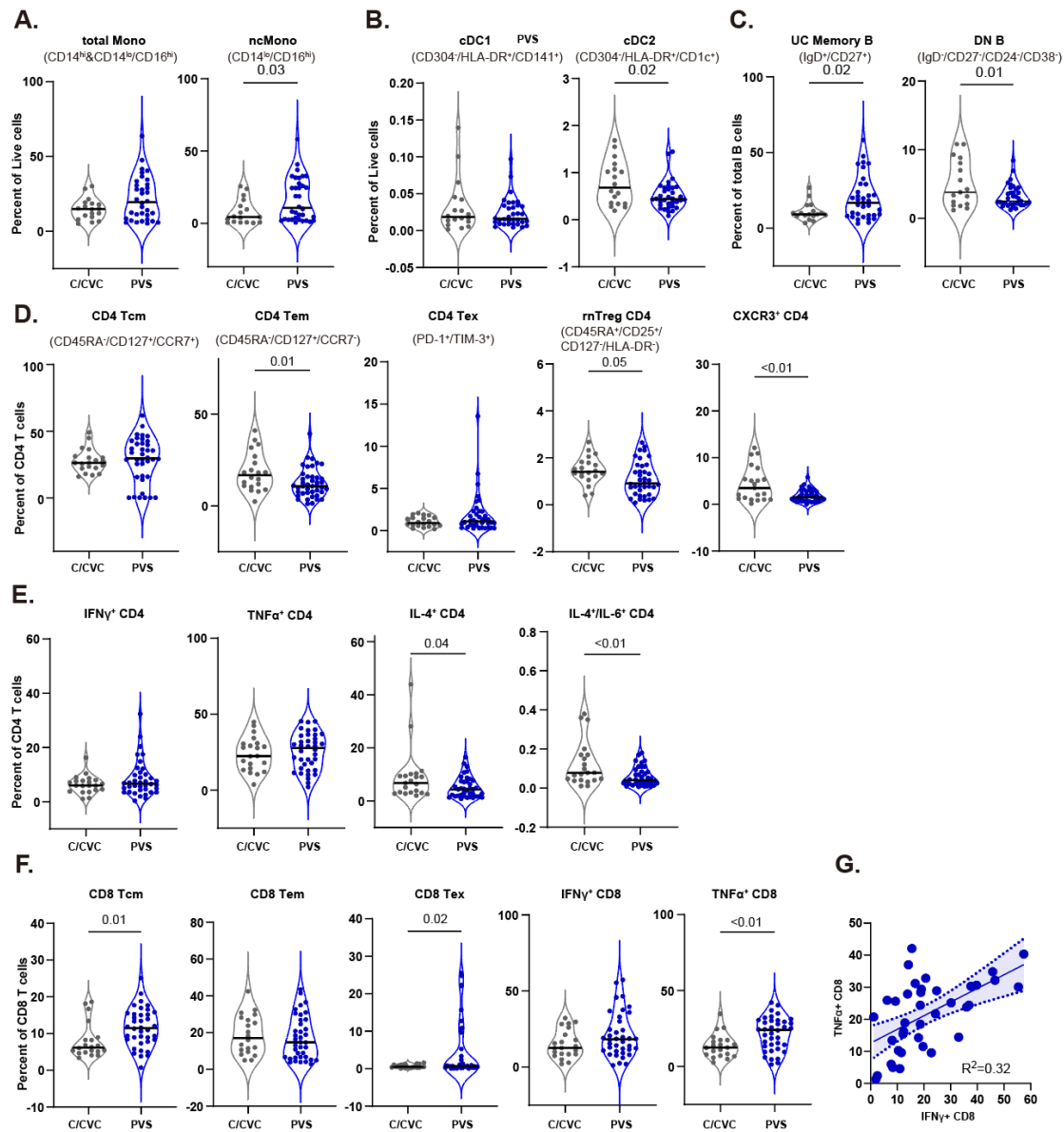
595 Table S5: Four group autoantibody analysis results

596 Table S6: Two group analysis results of soluble modulators

597 Table S7: Four group analysis results of soluble modulators

598 Table S8: External MY-LC cohort SPEAR S1 assay results

616



617

618 **Figure 2: Immune cell feature of myeloid and lymphoid cells in PVS patients. A-B.**

619 Violin plots of myeloid peripheral blood mononuclear populations (PBMCs) plotted by

620 groups as percentages of respective parent populations (live cells). **C.** Violin plots of B

621 lymphocyte subsets from PBMCs plotted as percentages of respective parent

622 populations (total B cells). **D.** Violin plots of various CD4 T cell subsets. **E.** Violin plots of

623 various cytokine-producing CD4 T cell subsets. **F.** Violin plots of various CD8 T cell

624 subsets and cytokine-producing CD8 T cell subsets. Significance differences were

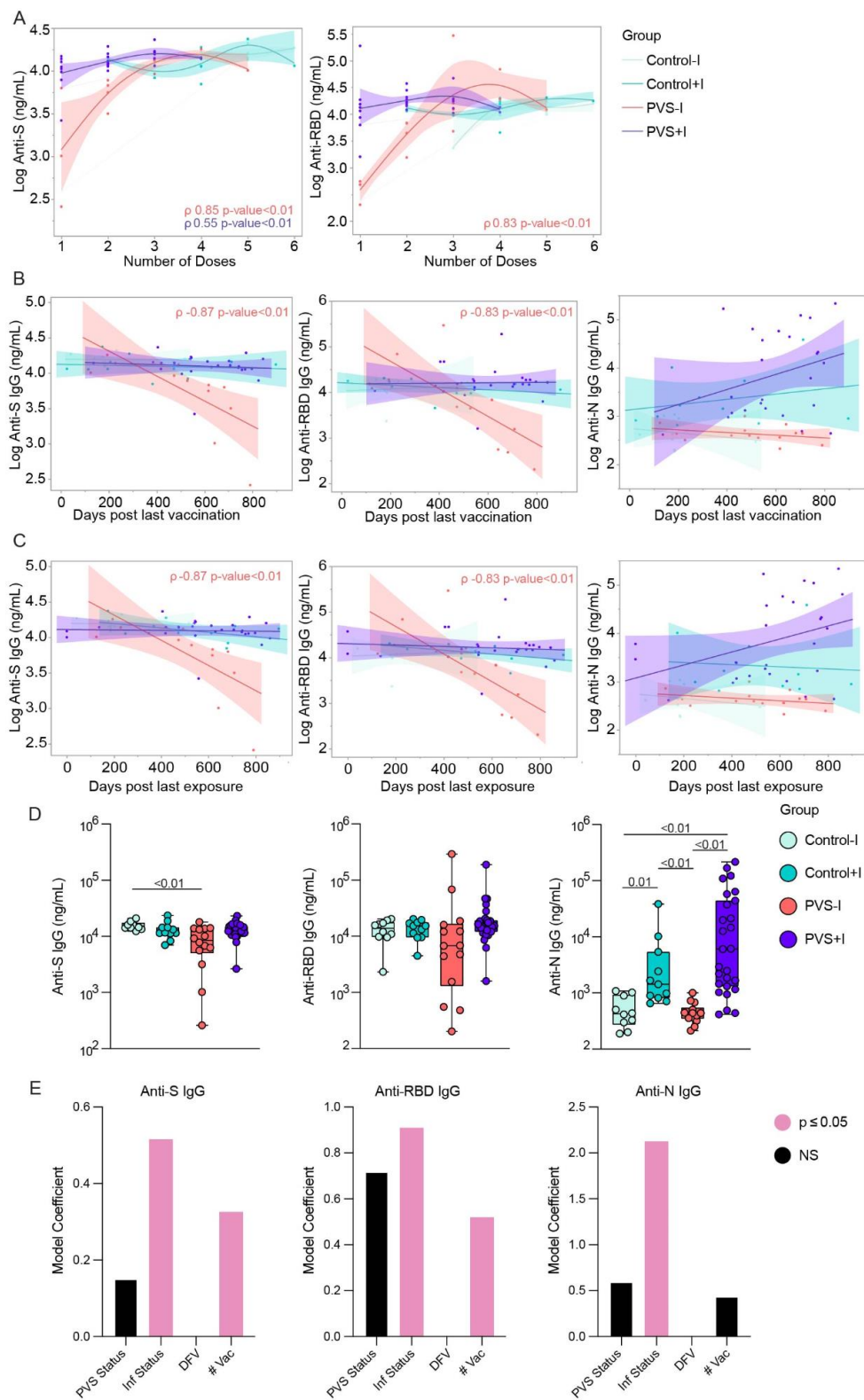
625 assessed using Mann-Whitney U tests with Benjamini–Hochberg false-discovery rate

626 (FDR) correction for multiple comparisons. **G.** Linear regressions of TNF α producing

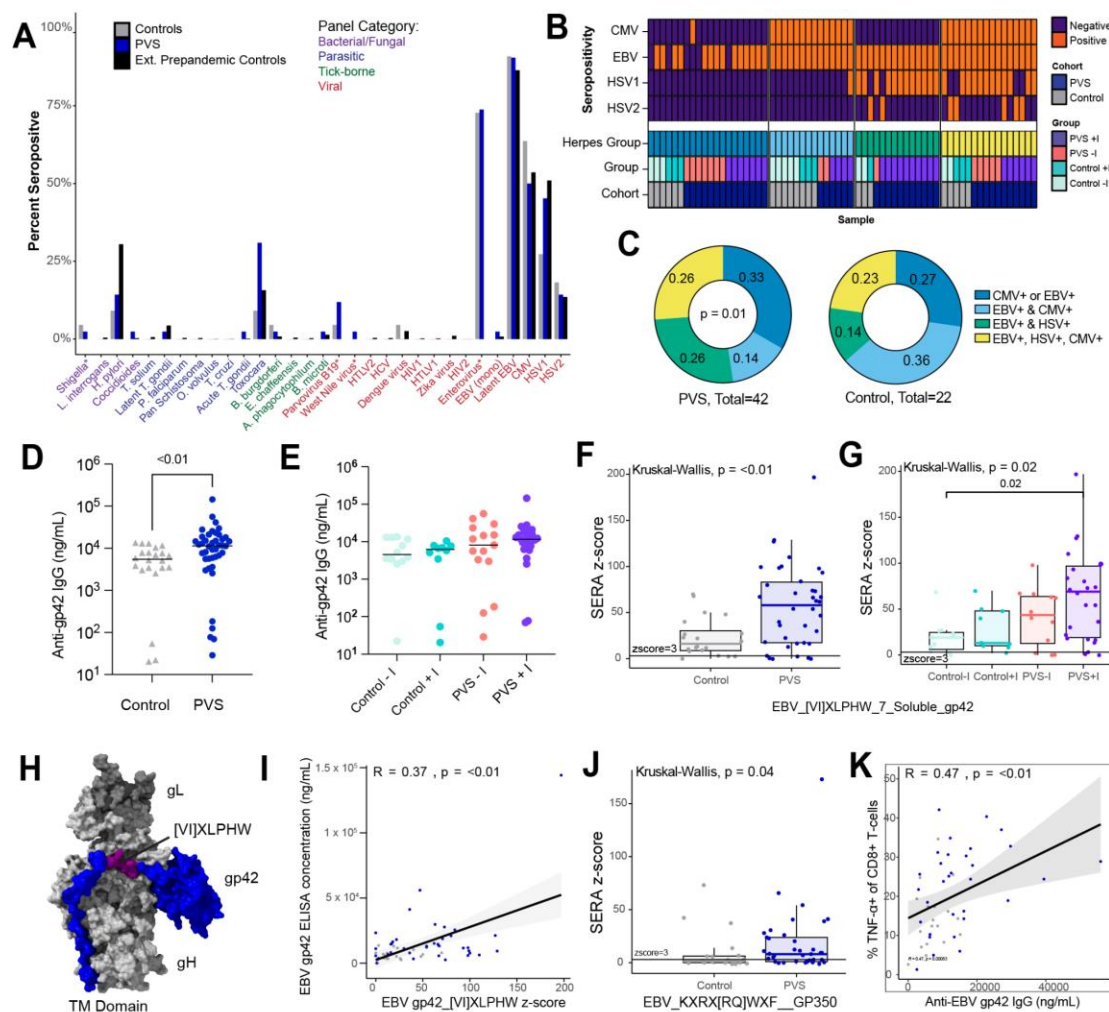
627 CD8 T cells and IFN γ producing CD8 T cells. Spearman's correlation was calculated

628 with corresponding p-values. Dotted lines depict linear regressions, with the area inside

629 representing 95% CI.



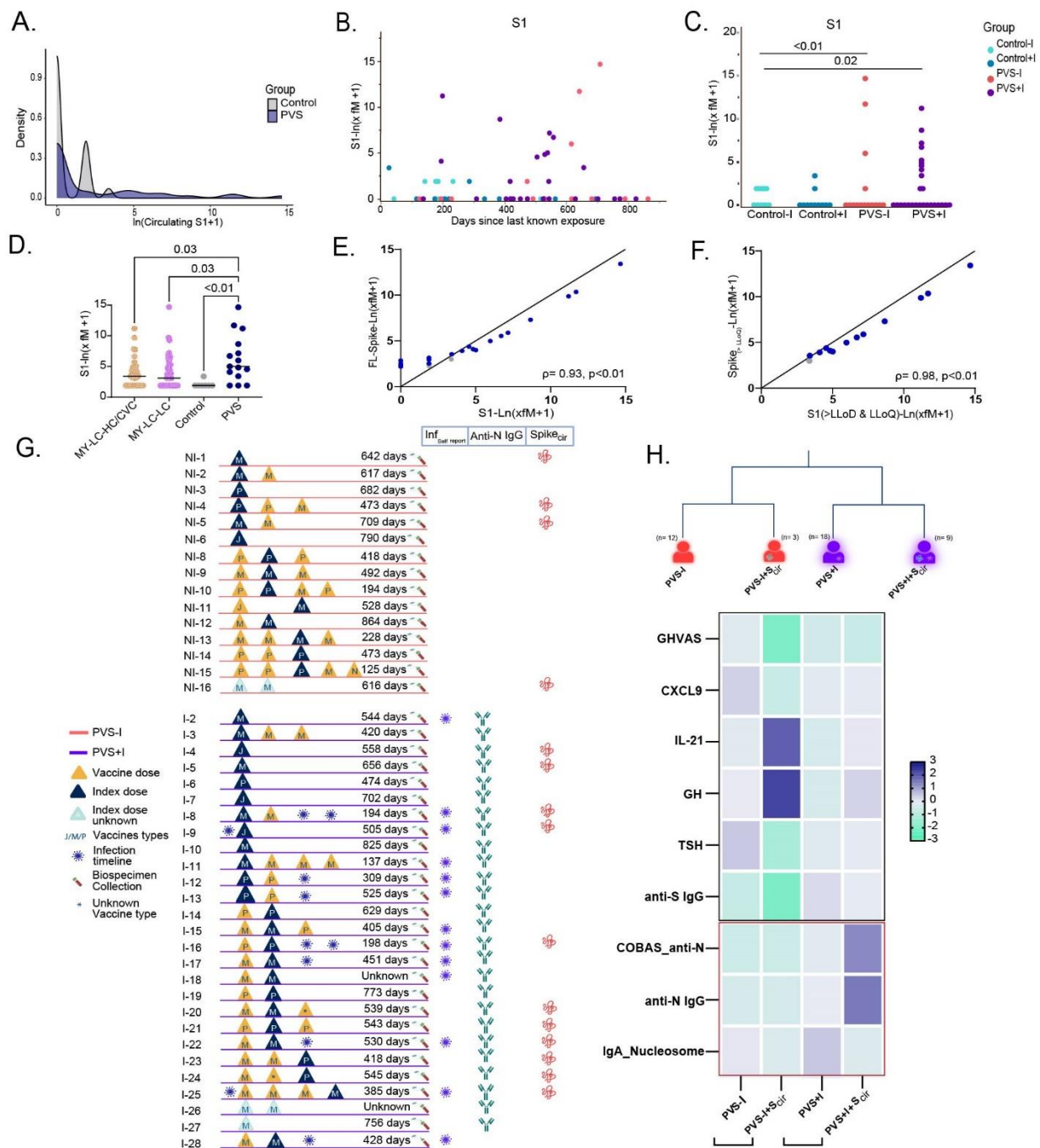
631 **Figure 3: Plasma reactivity to SARS-CoV-2 antigens. A.** Correlation comparisons of
632 virus-specific ancestral anti-S and anti-RBD IgG levels by number of COVID vaccine
633 doses. **B.** Correlation and linear regression comparisons of virus-specific ancestral anti-
634 S, anti-RBD and Anti-N IgG levels by days post last vaccination. **C.** Correlation and
635 linear regression comparisons of virus-specific ancestral anti-S, anti-RBD and Anti-N
636 IgG levels by days post last exposure. Regression lines are shown colored by groups
637 Control-I, Control+I, PVS-I, and PVS+I as indicated in the figure legend. Spearman's ρ
638 coefficients and linear regression significance are colored; accordingly, shading
639 represents 95% confidence interval. **D.** Plasma reactivity to ancestral S, RBD, and N
640 proteins measured by ELISA are shown by groups Control-I, Control+I, PVS-I, and
641 PVS+I. Significance of difference in group median values was assessed using Kruskal-
642 Wallis with Benjamini-Hochberg false-discovery rate (FDR) correction for multiple
643 comparisons. The central lines indicate the group median values, and the whiskers
644 show the 95% CI estimates. **E.** Generalized linear model analysis for virus-specific
645 ancestral anti-S, anti-RBD and Anti-N IgG levels. Model predictors are indicated on the
646 x axis and include days from vaccination (DFV) among others. Predictors with $p \leq 0.05$
647 are highlighted in pink to indicate significance, while non-significant predictors are
648 displayed in black. Detailed model results are shown in table S3.
649



650

651 **Figure 4: Elevated responses to Epstein Barr Virus in PVS patients.** **A.** Proportion
 652 of each group (PVS: $n = 42$, control: $n = 22$, pre-pandemic healthy control: $n = 3448$)
 653 seropositive for each of 31 common pathogen panels as determined by SERA, grouped
 654 by pathogen-type. Statistical significance determined by Fisher's exact test corrected
 655 with FDR (Benjamini Hochberg). Star indicates panels for which pre-pandemic healthy
 656 controls were not analyzed. **B.** Heatmap showing supervised clustering of SERA-
 657 determined seropositivity to EBV, CMV, HSV-1, and HSV-2 across samples. Clusters
 658 were named for their herpesvirus dominance and are labeled accordingly. **C.** Herpes
 659 seropositivity composition for each cohort. Significance of relative enrichment for each
 660 cluster was calculated using Chi-square test of observed composition vs. expected
 661 composition. **D, E** Plasma reactivity to EBV gp42 protein measured by ELISA shown by
 662 cohort, PVS and Control (**D**) and by groups Control-I, Control+I, PVS-I, and PVS+I (**E**).
 663 **F.** SERA-derived z scores for the gp42 motif [VI]XLPHW among EBV-seropositive
 664 individuals only, plotted by cohort, $n = 20$ (Control), $n = 38$ (PVS) (**F**) and group, $n = 11$
 665 (Control-I), $n = 9$ (Control+I), $n = 12$ (PVS-I), $n = 26$ (PVS+I) (**G**). The dashed line
 666 represents the z-score threshold for epitope positivity defined by SERA. **H.** Three-
 667 dimensional mapping of the PVS-enriched linear peptide sequence [VI]XLPHW (purple)
 668 onto EBV gp42 (blue) in a complex with gH (light grey) and gL (dark grey) (PDB: 5T1D).

669 **I.** Relationship between EBV gp42 [VI]XLPHW SERA z score and plasma concentration
670 of anti-gp42 IgG. **J.** SERA-derived z scores for the gp350 motif KXRX[RQ]WXF among
671 EBV-seropositive individuals only, plotted by cohort. The dashed line represents the z-
672 score threshold for epitope positivity defined by SERA. **K.** The relationship between
673 plasma concentration of IgG against EBV gp42 and the percentage of TNF α CD8+ T
674 cells (of total CD8+ T cells). For all box plots, the central lines indicate the group median
675 values, the top and bottom lines indicate the 75th and 25th percentiles, respectively, the
676 whiskers represent 1.5 \times the interquartile range. Each dot represents one individual.
677 Statistical significance of the difference in median values was determined using
678 Kruskal–Wallis tests with Post hoc Dunn’s test and Bonferroni–Holm’s method to adjust
679 for multiple comparisons. Correlation was assessed using Spearman’s correlation. The
680 black line shows linear regression, and shading shows the 95% CIs.
681

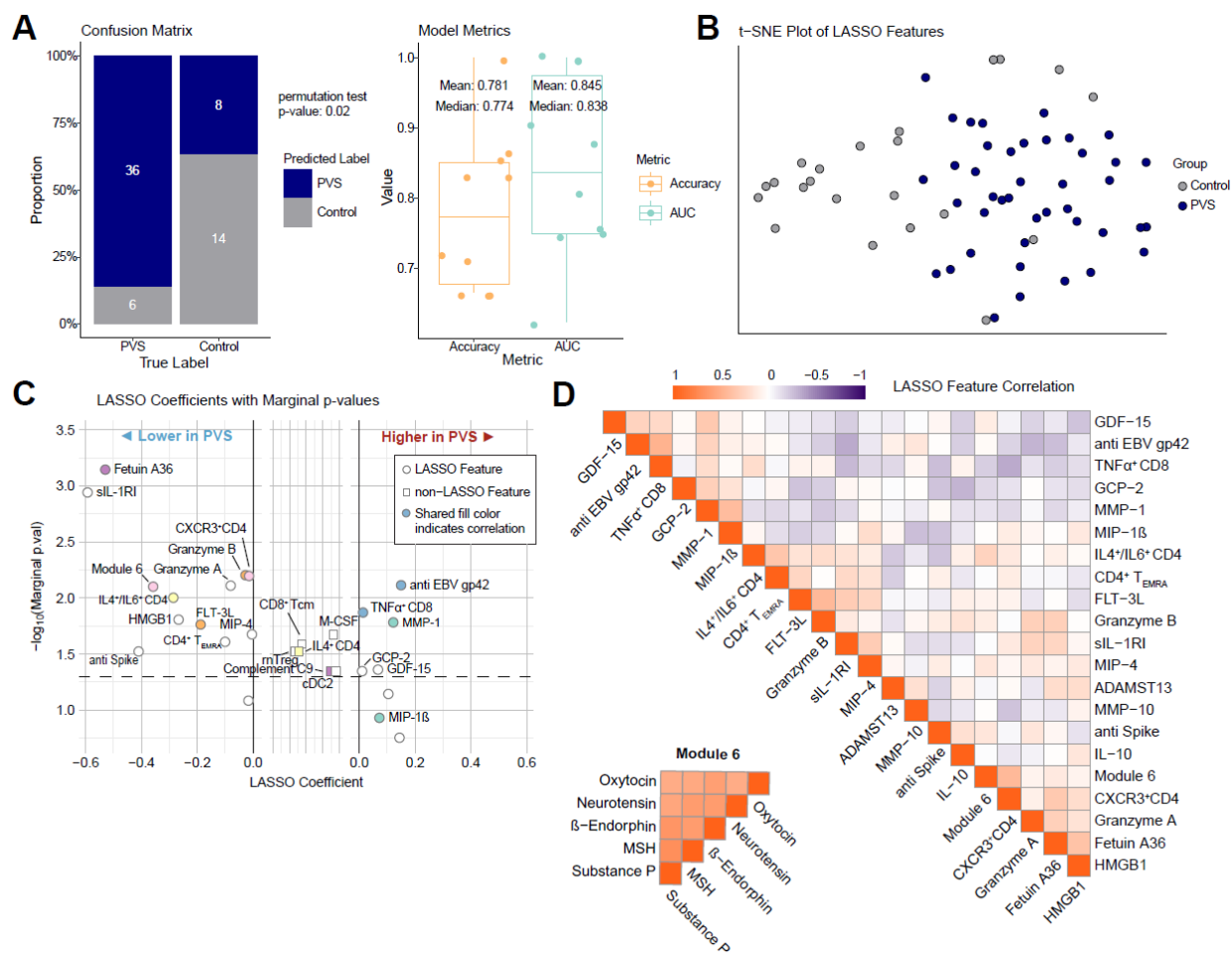


682
683
684
685
686
687
688
689

Figure 5: Circulating SARS-CoV-2 Spike protein. **A.** Density plots describing the distribution of circulating S1 levels across controls (n= 22) and PVS (n= 42) measured by SPEAR assays. **B.** Levels of circulating S1 in plasma days post last known self-reported exposure. **C.** Circulating S1 levels measured by SPEAR assay are shown across groups Control-I, Control+I, PVS-I, and PVS+I. A parametric test incorporating a zero-inflated Poisson model was used to account for the excess zeros in the data. **D.** Circulating S1 antigen levels above LLoD across cohort groups, MY-LC-HC/CVC (n=

690 41), MY-LC-LC (n=45), Control (n= 7), PVS (n= 15). **E.** Correlation between circulating
691 Spike protein assays using antibodies for the S1 subdomains and S1& S2 subdomains
692 (full length Spike). **F.** Correlation between circulating Spike protein assays among
693 samples with values above LLoD and LLoQ. Correlations were assessed using
694 Spearman's rank correlation. **G.** A participant-specific graphic representation illustrating
695 key events including vaccination, infection, sample collection and the presence or
696 absence of anti-N antibodies and circulation spike protein. Each participant is
697 represented by a single horizontal line and each vaccination event is marked by
698 triangles, and the index doses are highlighted in blue. The number of days between
699 latest exposure and biospecimen collection is also indicated. The abbreviations for the
700 vaccine types are: J for Jcovden (Johnson & Johnson), M for Spikevax (Moderna), and
701 P for Comirnaty (Pfizer-BioNTech). **H.** A classification tree of PVS participants based on
702 infection status with or without detectable S1 in circulation and a heatmap of distinct
703 demographic and immunological variables that differentiate PVS within the infected and
704 uninfected subgroups based on Mann-Whitney U tests.
705
706

707



708

709

710 **Figure 6: Machine Learning Results and Prominent Feature Identification. A.**

711 Confusion matrix inspired barplot describing actual and predicted labels for PVS and

712 control with a classification threshold of 0.65. The p-value from a permutation test is

713 also displayed to motivate model legitimacy. Boxplot displaying accuracy and AUC for

714 each outer fold of nested-cross validation with both inner and outer loops set to 10.

715 Mean and median accuracy and AUC stats are also shown. **B.** t-SNE plot showing a

716 two-dimensional similarity-based representation of the LASSO variable space. **C.**

717 Scatter plot containing all variables identified through either LASSO or marginal logistic

718 regression models. Features only identified by marginal test, and not by LASSO are

719 displayed in the center panel with coefficient of 0. Points with a shared color display

720 Pearson correlation > 0.4. **D.** Correlation heatmaps of LASSO-selected and WGCNA

721 module six features, with features clustered through hierarchical clustering.

722

723 **MATERIALS AND METHODS**

724 **Ethics Statement and Study Design**

725 A retrospective, decentralized, exploratory case-control study was conducted on
726 vaccinated participants ≥ 18 years of age enrolled in the LISTEN study which recruited
727 individuals from the Hugo Health (Guilford, CT, USA) Kindred community⁵¹. This study
728 was approved by the Yale University Institutional Review Board on April 1, 2022 (HIC#
729 2000032207) and conforms to the Declaration of Helsinki and STROBE reporting
730 guidelines^{52,53}. Informed consent was provided by participants electronically. Only those
731 residing in contiguous US states were included in this arm of the study. Each participant
732 was assigned a unique identifier as part of the de-identification protocol managed by the
733 study coordinator. These identifiers were kept confidential and were not accessible to
734 anyone outside the research team.

735 A subset of the MY-LC cohort recruited from within the Mount Sinai Healthcare System
736 and the Centre for Post COVID Care at Mount Sinai Hospital which included healthy
737 and convalescent controls and long COVID participants were included for validation of
738 one of the assays. The biospecimens were collected between 2021 and 2023 as
739 approved by the Mount Sinai Program for the Protection of Human Subjects (#20-
740 01758). The recruitment for the MY-LC cohort has been elaborately described in the
741 Klein et al study²².

742 **Inclusion and Exclusion Criteria**

743 Given the exploratory nature of this study, no power analysis was done. Only US
744 residents were included. Cases were defined as participants who self-reported having
745 PVS after administration of the primary or booster series of COVID-19 vaccine and
746 controls were vaccinated individuals who reported no adverse events. The cohort
747 consisted of both individuals with self-reported history of SARS-CoV-2 infection or
748 otherwise. Participants with a medical history of arthritis, asthma, autoimmune disease,
749 blood clots, cancer, cerebrovascular conditions, chronic lung disease (including
750 emphysema, chronic bronchitis, chronic obstructive pulmonary disease), cystic fibrosis,
751 diabetes, Ehlers Danlos Syndrome, myocardial infarction, heart conditions, heart failure,
752 history of organ transplant, immunocompromised state, kidney disease, liver disease,
753 Lyme disease, mast cell activation syndrome, myalgic encephalomyelitis/chronic fatigue
754 syndrome, neurologic conditions, postural orthostatic hypotension, spinal disorder(s),
755 tremors/Internal vibrations, hyperthyroidism, hormonal migraines, schizophrenia
756 spectrum disorders, bipolar disorders and somatic symptoms before 2020 were
757 excluded. Participants who self-reported having both long COVID and PVS were also
758 excluded from this study. Among the selected cases, those with GHVAS scores of ≤ 75
759 at the time of screening were included, while among the controls, only those with
760 GHVAS scores of ≥ 80 or general health of very good/excellent were included.

761 Electronic health records, shared by participants via Hugo Health were accessed to
762 obtain a participant-wise list of ongoing medications to exclude participants on
763 immunosuppressants.

764 **Survey Instruments**

765 Members of the Hugo Health Kindred community were offered a set of specialized
766 baseline survey instruments. The first questionnaire was geared towards gathering
767 information on demographic variables such as age, biological sex, ethnicity as well as
768 self-reported medical conditions such as long COVID and PVS. The subsequent
769 surveys were on self-reported SARS-CoV-2 infection, vaccination profiles and medical
770 history, comorbidities, PVS symptoms¹⁰. These survey responses were analyzed, and
771 a subset of the participants were included in this arm of the study based on the inclusion
772 and exclusion criteria mentioned above.

773 At the time of biospecimen collection, the participants were asked to rate their general
774 health status using the following options: excellent, very good, good, fair, poor
775 and do not know and a general health visual analogue scale (GHVAS) having a
776 numerical value between 0 to 100, where higher scores represented better health.
777 These responses were recorded by ExamOne phlebotomists on LISTEN-ExamOne
778 requisition forms and shipped with biospecimens. Participants also completed the
779 Patient-Reported Outcomes Measurement Information System-29 version 2
780 questionnaire (PROMIS-29 v2) through the Hugo Health interface within 24 hours of
781 biospecimen collection.

782 **Biospecimen collection**

783 Whole blood samples were collected in lithium-heparin-coated vacutainers (BD 367880,
784 BD Biosciences) from participants' homes or agreed upon locations, by ExamOne
785 phlebotomists (Quest Diagnostics). Biospecimens were shipped overnight at regulated
786 temperatures to Yale University in New Haven, CT on the day of collection. Upon
787 receipt, collection tubes were de-identified according to protocol and study identifiers
788 were provided. Samples were processed thereafter and within 48 hours of collection.

789 Biorender⁵⁴ was used to create a graphical schematic on study design, cohorts and
790 assays.

791 **Plasma and PBMC isolation**

792 Plasma and peripheral blood mononuclear cells (PBMC) were isolated following
793 protocols published previously²². Briefly, whole blood samples were centrifuged at 650g
794 for 10 minutes at room temperature without brake. Plasma from each participant were
795 pooled together to a single 15 mL polypropylene conical tube, aliquoted and stored at -
796 80°C. PBMCs were subsequently isolated using SepMate tubes (StemCell) following
797 manufacturer's instructions. Freshly isolated PBMCs were used for downstream flow
798 cytometry analyses or were gradually cryopreserved using a freezing container and
799 then transferred to liquid nitrogen tanks for long-term storage.

800 **Anti-SARS-CoV-2 nucleocapsid (N) antibody testing**

801
802

803 The EUA-cleared Elecsys Anti-SARS-CoV-2 immunoassay (Roche Diagnostics,
804 Indianapolis) was used to independently determine prior history of SARS-CoV-2
805 infection among participants at the Yale-New Haven hospital clinical research
806 laboratory. Plasma aliquots were analyzed using this double-antigen sandwich (DAGS)
807 immunoassay that measures total high affinity antibodies directed against the viral N
808 protein. The COBAS e801 platform (Roche Diagnostics, Switzerland) was used
809 according to the manufacturer's instructions as validated earlier²⁵.

810
811 Both self-reported infection history and the results of this assay were used to determine
812 a participant's history of SARS-CoV-2 infection. Participants were classified as infection
813 positive if they reported a past infection but had no detectable anti-N antibodies, or if
814 they had detectable anti-N antibodies regardless of not reporting a prior infection.

815 **Flow cytometry**

816 Flow cytometry was performed on an Attune NxT Flow Cytometer (Thermo Fisher)
817 using NxT v 5.31.0 software as described previously²². In brief, 1 to 2 million freshly
818 isolated PBMCs were plated in round-bottom 96 well plates for surface and intracellular
819 staining. Dead cells were stained using Live/Dead Fixable Aqua (ThermoFisher) on ice
820 followed by a wash and Fc receptor blocking (Human TruStain FcX™, Biolegend). A
821 total of 30 fluorophore conjugated antibodies were used in three different combinations
822 for surface staining to define myeloid, B and T cell lineages. Quantitation of intracellular
823 cytokine staining with stimulation was executed by incubating cells in 1x cell stimulation
824 cocktail (eBioscience) without protein transport inhibitor for 1 hour in 10% FBS cRPMI
825 followed by 4 hour incubation in 1x stimulation cocktail with protein transport inhibitor(
826 eBioscience) in 10% FBS cRPMI at 37°C. Cells were permeabilized using 1x
827 permeabilization buffer from the FOXP3/Transcription Factor Staining Buffer Set
828 (eBioscience) at 4 °C and stained after Fc receptor blocking. FlowJo v.10.8 (BD) was
829 used for data analysis. All the markers and gating strategies were as described
830 previously and noted in the Flow Repository (ID: FR-FCM-Z8FZ) ²².

831 832 **SARS-CoV-2 antibody specific immunoassays**

833
834 Enzyme-linked immunosorbent assays (ELISA) were run as described earlier ⁵⁵. Briefly,
835 96-well MaxiSorp plates (Thermo Fisher Scientific, #442404) were coated at a
836 concentration of 2 µg/ml in PBS with recombinant SARS-CoV-2 S protein (50 µl per
837 well; ACROBiosystems, #SPN-C52H9-100 µg) or RBD (ACROBiosystems, #SPD-
838 C52H3-100 µg) or nucleocapsid protein (NUN-C5227-100 µg, ACROBiosystems) and
839 incubated overnight at 4°C. After removing the coating buffer, the plates were blocked
840 with 200 µl of blocking solution (PBS with 0.1% Tween-20 and 3% milk powder) for 1
841 hour at room temperature (RT). Plasma samples were diluted 1:1500 (for Anti-Spike
842 and Anti-RBD) and 1:400 (for Anti-N) in buffer (PBS with 0.1% Tween-20 and 1% milk
843 powder), and 100 µl of the diluted plasma were added to the wells for 2 hours at RT. A
844 standard curve was generated using serially diluted Human anti-Spike [SARS-CoV-2
845 Human Anti-Spike (AM006415) (Biolegend, #938602)] and anti-nucleocapsid SARS-
846 CoV-2 human anti-nucleocapsid (1A6) (ThermoFisher Scientific, #MA5-35941) . After

847 three washes with PBS-T (PBS with 0.1% Tween-20), 50 μ l of horseradish peroxidase-
848 conjugated anti-Human IgG antibody (GenScript #A00166; 1:5000) diluted in buffer was
849 added to each well and incubated for an hour. The plates were then developed with 100
850 μ l of TMB (3,3',5,5'-tetramethylbenzidine) Substrate Reagent Set (BD Biosciences
851 #555214) and read at wavelengths of 450 and 570 nm.

852

853 **Linear peptide profiling & ELISA validation**

854

855 Plasma aliquots were shipped to Serimmune on dry ice and the SERA assays were
856 used to discover pathogen specific immunoglobulin G responses to linear peptides as
857 described earlier^{22,56}. The published PIWAS method⁵⁷ was used to identify candidate
858 SARS-CoV-2 antigens and epitopes from the UniProt reference proteome
859 (UP000464024). Briefly, the 12mer peptide sequences captured for each sample were
860 broken down into 5- and 6-mers (k-mers) and tiled onto the proteome. An enrichment
861 for each k-mer was then calculated as described⁵⁷ based on the frequency of that kmer
862 in the dataset relative to the frequency expected by random chance. Outliers were
863 identified by comparing the enrichment to a large pre-pandemic cohort ($n = 1,500$). An
864 outlier threshold and outlier sum statistic were then derived, and a p value was
865 calculated by comparing the observed outlier sum statistic to the null distribution. Outlier
866 sum threshold was set to the 99.5th percentile value of all positions with FDR-adjusted
867 $p > 0.001$. All sequence positions that exceeded both thresholds were identified, and
868 adjacent positions were merged into regions representing candidate epitopes on the
869 protein.

870

871 **ELISA Validation**

872 96-well MaxiSorp plates (Thermo Scientific #442404) were coated with 200 ng per well
873 of recombinant EBV gp42/ protein (MyBioSource #MBS430545) in PBS and incubated
874 overnight at 4 °C. Plates were emptied and incubated with 3% Omniblock non-fat dry
875 milk (American Bioanalytical #AB10109-00100) in PBS with 0.1% Tween 20 (Sigma-
876 Aldrich, #P7949) for 2 hours at RT and then overnight at 4°C. Plates were washed 3x
877 with 200 μ l wash buffer (PBS 0.1% Tween 20). Samples and a serial dilution curve of
878 monoclonal antibody against EBV gH/gL/gp42 (MyBioSource #MBS430548) were
879 diluted in 1% Omniblock non-fat dry milk in PBS and added to the plate to incubate for 1
880 hour at RT. Plates were washed 3x with a wash buffer. Goat anti-human IgG Fc HRP
881 (Sigma Aldrich, #AP112P) diluted 1:5000 or goat anti-mouse IgG Fc HRP (Southern
882 Biotech #1030-05) diluted 1:6000 in 1% Omniblock non-fat dry milk in PBS was added
883 to the plates and incubated for 50 minutes at RT. Plates were washed 5x. Plates were
884 developed with 50 μ l of TMB Substrate Reagent Set (BD Biosciences #555214) and the
885 reaction was stopped after 5 min by the addition of 2 N sulfuric acid. Plates were then
886 read at a wavelength of 450 nm.

887 **Immunoglobulin Isotyping & Autoantigen arrays**

888 Plasma aliquots were shipped to Eve Technologies on dry ice. Two assays were run to
889 determine the concentrations of the different Ig subtypes and IgG isotypes in the

890 plasma, namely, Immunoglobulin Isotyping 6-Plex Custom Assay (IgG1, IgG2, IgG3,
891 IgG4, IgA, and IgM) and the Human-IgE-1-Plex-Custom-Assay.

892 Multiplex IgM and IgG autoantibody reactivities were analyzed at the Microarray Core
893 facility at the University of Texas Southwestern Medical Center, Dallas, TX, USA as
894 described previously^{58,59}. Plasma aliquots were shipped to the facility on dry ice. A
895 total of 120 human antigens, autoantibodies against which have been associated with
896 various immune-related diseases or allergic disorders were included, namely, ACE2,
897 Aggrecan, Albumin, α Fodrin, Amyloid β (1-40), Amyloid β (1-42), AQP4, BAFF,
898 BCOADC-E2, BPI, Calprotectin/S100, CD4, CD40, CENP-A, CENP-B, collagen I,
899 collagen II, collagen III, collagen IV, collagen V, Complement C1q, Complement C3,
900 Complement C4, Complement C5, Complement C6, Complement C7, Complement C8,
901 Complement C9, CRP, Cytochrome C, DFS70, dsDNA, EJ, Factor B, Factor H, Factor I,
902 Factor P, fibrinogen Type I-S, fibronectin, GAD65, GBM, genomic DNA, Gliadin, gp210,
903 GP2, H/K-ATPase, histone, histone H1, histone H2A, histone H2B, histone H3, HSPG,
904 IA-2, IF, IFN γ , IL-6, IL-12/NKSF, IL-17A, Jo-1, KS, KU (P70/P80), La/SS-B, Laminin,
905 LC1, LKM 1, LPS, lysozyme, M2, MBP, MDA5, Mi-2, mitochondrion, MPO, myosin,
906 Nrp1, nucleolin, nucleosome, Nup62, NXP2, OGDC-E2, P0, P1, P2, PCNA, PDC-E2,
907 PL-7, PL-12, PM/ScI-75, PM/ScI 100, PR3, proteoglycan, prothrombin, Ro/SS-A(52
908 kDa), Ro/SS-A(60 Kda), SAE1/SAE2, ScI-70, SLA/LP, Sm, Sm/RNP, SmD, SmD1,
909 SmD2, SmD3, SP100, SRP54, ssDNA, Tau, thyroglobulin, TIF1 γ , TLR4, TNF α , TPO,
910 tTG, U1-snRNP 68/70kDa, U1-snRNP A, U1-snRNP C, U-snRNP B/B', Vimentin,
911 Vitronectin and β 2-Glycoprotein 1) were printed on 16 pad nitrocellulose Fast Analyte
912 Scanning Technology (FAST) slides. Eight positive control proteins (human IgM and
913 IgG, anti-human IgM and IgG) were also imprinted on the arrays as positive controls.
914 Briefly, DNase I (Thermo Fisher, #AM2222) treated plasma samples were applied onto
915 the autoantigen arrays at 1:50 dilution. Binding was detected using cy5-labeled anti-
916 human IgM and cy3-labeled anti-human IgG antibodies (Jackson ImmunoResearch
917 Laboratories, #IgM 109-606-129, #IgG-109-166-098), and the array slides were
918 scanned at wavelengths of 532 nm for and 635 nm for cy5 and cy3 respectively using a
919 Genepix 4400A scanner. Genepix Pro v7.0 software (Molecular Devices) was used to
920 analyze the image and generate the genepix report (GPR) files. The net signal intensity
921 (NSI) values for each antigen were calculated by subtracting the negative control
922 values. Normalized NSI values calculated using built-in positive controls were used for
923 case-control analyses. Signal-to-noise ratios (SNR) for each antigen were calculated
924 using the formula: (foreground median value – background median value)/standard
925 deviation. Antigens with SNR<3 in more than 10% of all samples were excluded. Log₂
926 fold-change (FC) values between cases and controls were calculated.

927 **Multi-target soluble plasma factor analysis**

928 Plasma aliquots stored at –80 °C were shipped to Eve Technologies on dry ice. All the
929 samples were run in the same batch to avoid batch effects. The analytes covered by the
930 following panels were assayed: Human Cytokine/chemokine 96-Plex Discovery Assay
931 (HD96), Human Cytokine/chemokine Panel 4 12-Plex Discovery Assay(HDIV12),
932 Human Adipokine 5-Plex Discovery Assay (HDADK5), Human Cardiovascular Disease
933 Panels 2 and 3 (HDCVD8, HDCVD9), Human Complement Expanded Panel 1 9-Plex

934 Discovery Assay (HDCMPEX1), Human MMP and TIMP 13-Plex Discovery Assay
935 (HMMP/TIMP-S,P), Human Neuropeptide 5-Plex assay (HNPMAG-35K), Human
936 Soluble Cytokine Receptor 14-Plex Discovery Assay (HDSCR14), Multi-Species TGF- β
937 3-Plex Discovery Assay (TGF β 1-3), Human Pituitary Panel 1 7-plex (HPTP1),
938 Multispecies Hormone Panel 5-Plex (MSHMAG-21K) and Arginase-1.

939 To quantitate the total plasma testosterone levels, competitive ELISAs were performed
940 as per manufacturer's recommendations with the exception of the dissociation time
941 which was increased to 30 mins to ensure complete release of protein bound
942 testosterone (Thermo Fisher, #EIATES). Only analytes with less than 20% missing
943 values were included in the analyses.

944 **SPEAR SARS-CoV-2 Spike Protein Immunoassay**

945 The SARS-CoV-2 S1 antigen was measured in plasma samples using the SPEAR
946 SARS-CoV-2 Spike Protein Immunoassay, developed and performed by Spear Bio, Inc.
947 (Woburn, MA). The SPEAR immunoassay employs two specially designed DNA probes
948 conjugated to target-specific antibodies that undergo a two-step successive polymerase
949 extension reaction when bound to the target analyte, producing a unique DNA strand for
950 analyte quantification with extremely low background. The assay uses antibodies
951 specific to two distinct epitopes within the S1 subunit (GenScript, #A02052, #A02058).
952 Plasma samples first underwent a reduction treatment to release S1 from potentially
953 bound-blocking antibodies. In brief, plasma samples were incubated with SPEAR-
954 specific reducing reagent for 15 minutes at 37°C, followed by the addition of a SPEAR-
955 specific reduction quenching solution before incubation with the probes. The treated
956 samples were then diluted tenfold in the assay diluent and incubated with S1-specific
957 antibody-oligonucleotide probes for 2 hours at 37°C. This was followed by a 30-minute
958 reaction to convert the presence of the analyte into DNA signal strands. The DNA signal
959 strands were quantified through qPCR using the QuantStudio 12K Flex (ThermoFisher)
960 with primers and qPCR probes specific to the DNA signal strand sequence.

961 Standard curves were generated from a stock solution S1 spike protein spanning from
962 0.0625 – 100,000 fM. Results from standard curves were used to generate sigmoidal
963 four parameter logistic (4PL) fits in GraphPad Prism (v 10) software. Sample
964 concentration results were calculated from run specific 4PL fit, multiplied by dilution
965 factor. Lower limit of detection (LLoD) was calculated as the concentration
966 corresponding to 2.5 standard deviation + mean zero in signal (18 replicates per run).
967 Lower limit of quantification (LLoQ) was derived from the sample precision profile
968 [measured concentration vs. concentration coefficient of variation (CV)] as the
969 measured concentration at which a sample precision fit intersects at 20% concentration
970 CV. The LLoQ and the LLoD values were calculated to be 20.5 fM and 5.64 fM as a
971 mean of 6 runs. Values below LLoQ were replaced by the assay LLoD value and all
972 values below LLoD were imputed as zero. The values were then transformed using the
973 natural logarithm ($\ln(x + 1)$) for comparisons. Analysis between groups was performed
974 through a generalized regression model for zero inflation distribution (ZI Poisson)
975 applying Tukey's HSD to each pair.

976 To further validate the S1 assays, independent SPEAR immunoassays were performed
977 to detect the full-length protein. This assay utilized oligonucleotide probes conjugated to
978 antibodies targeting both the S1 (GenScript, #A02052) and S2 (R&D Systems, #
979 MAB11362-100) subdomains. The S1 subunit binding antibody was common to both
980 the assays.

981 **General Statistical Analyses**

982
983 The association of variables between case and control cohorts and among the four
984 groups were assessed using Fisher's exact test for categorical variables. For pair-wise
985 comparisons, Mann-Whitney tests with Benjamini-Hochberg correction for multiple
986 testing were used. For four-group comparisons, Kruskal-Wallis tests followed by
987 Conover's post-hoc test with Benjamini-Hochberg method for multiple comparisons
988 correction were applied to continuous variables. Post hoc tests were also performed
989 using Dunn's test with Bonferroni-Holm's method to adjust for multiple comparisons.
990 Both one and two-sided p values ≤ 0.05 were considered statistically significant.
991 Spearman's rank correlation tests were utilized to analyze the relationships between cell
992 populations, antibody levels, anti-SARS-CoV-2 antibody levels, the number of
993 vaccinations, days post-vaccination, days post-exposure and in assay validations.

994 **Linear modeling**

995
996
997 Generalized linear modeling was performed linear regression with outputs of model
998 coefficients for each variable and associated p-values ≤ 0.05 were considered
999 statistically significant. All analysis was performed using R. Standard generalized linear
1000 model function listed is below.

$$1001 \quad y \sim 1 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots$$

1002 **Machine Learning Analysis**

1003
1004
1005 To predict PVS and control outcomes, we utilized a two-step approach involving
1006 weighted gene co-expression network analysis (WGCNA) and the construction of a
1007 generalized LASSO linear model^{31,60}. Initially, we preprocessed the data via quantile
1008 normalization to address outliers and skewness, with average ranks assigned in case of
1009 ties. Due to the presence of highly correlated features, we applied WGCNA to identify
1010 clusters of features with similar expression patterns. Specifically, we set the soft
1011 thresholding power parameter to 7, which is crucial for transforming the raw correlations
1012 into adjacency values, by maintaining the highest mean connectivity and achieving a
1013 model fit higher than 0.8. Using this selected parameter, we computed the network
1014 adjacency based on Pearson correlation within a signed network. We then constructed a
1015 topological overlap matrix (TOM) and converted it into its corresponding dissimilarity
1016 measure. Hierarchical clustering was performed on this measure with ward.D2 linkage,
1017 and the dendrogram was pruned to enforce a minimum cluster size of 5. This process
1018 allowed us to identify 6 distinct modules.

1019
1020 We built the LASSO model using raw features from module 1, a large module with
1021 numerous features that was not further clustered, along with the eigengenes (first

1022 principal component) of the other 5 modules. To compensate for small sample size and
1023 assess model performance, we conducted nested cross-validation with both inner and
1024 outer loops set to 10, and adjusted the classification threshold to 0.65 to combat class
1025 imbalance. Additionally, a permutation test was performed by shuffling the PVS label
1026 within the infection and non-infection group, demonstrating the fitted model performed
1027 better than random chance. We then refitted the LASSO model using all available data
1028 and applied the min criterion rule to select important features. Given that LASSO does
1029 not inherently provide p-values due to its regularization nature, we also assessed the
1030 marginal significance of each individual feature using logistic regression. To better
1031 understand model performance, t-SNE, or t-distributed Stochastic Neighbor Embedding,
1032 was used to create low-dimension plots of the LASSO variable space.

1033
1034 Lastly, to ensure model robustness, a secondary workflow was utilized to examine
1035 variable importance. For this, a random forest model was created utilizing nested cross-
1036 validation and SMOTE up-sampling to compensate for class imbalance. Random forest
1037 permutation importance was then used as a metric to examine the top features.

1038
1039
1040

1041 References

1042

- 1043 1. Watson, O.J., Barnsley, G., Toor, J., Hogan, A.B., Winskill, P., and Ghani, A.C. (2022).
1044 Global impact of the first year of COVID-19 vaccination: a mathematical modelling study.
1045 *Lancet Infect Dis* 22, 1293-1302. [10.1016/S1473-3099\(22\)00320-6](https://doi.org/10.1016/S1473-3099(22)00320-6).
- 1046 2. Zheng, C., Shao, W., Chen, X., Zhang, B., Wang, G., and Zhang, W. (2022). Real-world
1047 effectiveness of COVID-19 vaccines: a literature review and meta-analysis. *Int J Infect*
1048 *Dis* 114, 252-260. [10.1016/j.ijid.2021.11.009](https://doi.org/10.1016/j.ijid.2021.11.009).
- 1049 3. Xie, Y., Choi, T., and Al-Aly, Z. (2024). Postacute Sequelae of SARS-CoV-2 Infection in
1050 the Pre-Delta, Delta, and Omicron Eras. *N Engl J Med* 391, 515-525.
1051 [10.1056/NEJMoa2403211](https://doi.org/10.1056/NEJMoa2403211).
- 1052 4. De Domenico, M. (2023). Prevalence of long COVID decreases for increasing COVID-19
1053 vaccine uptake. *PLOS Glob Public Health* 3, e0001917. [10.1371/journal.pgph.0001917](https://doi.org/10.1371/journal.pgph.0001917).
- 1054 5. (2024). In *Evidence Review of the Adverse Effects of COVID-19 Vaccination and*
1055 *Intramuscular Vaccine Administration*, D. Rosenberg, O.K. Kumova, K. Stratton, and
1056 A.R. Bass, eds. [10.17226/27746](https://doi.org/10.17226/27746).
- 1057 6. Faksova, K., Walsh, D., Jiang, Y., Griffin, J., Phillips, A., Gentile, A., Kwong, J.C.,
1058 Macartney, K., Naus, M., Grange, Z., et al. (2024). COVID-19 vaccines and adverse
1059 events of special interest: A multinational Global Vaccine Data Network (GVDN) cohort
1060 study of 99 million vaccinated individuals. *Vaccine* 42, 2200-2211.
1061 [10.1016/j.vaccine.2024.01.100](https://doi.org/10.1016/j.vaccine.2024.01.100).
- 1062 7. Schultz, N.H., Sorvoll, I.H., Michelsen, A.E., Munthe, L.A., Lund-Johansen, F., Ahlen,
1063 M.T., Wiedmann, M., Aamodt, A.H., Skattor, T.H., Tjonnfjord, G.E., and Holme, P.A.
1064 (2021). Thrombosis and Thrombocytopenia after ChAdOx1 nCoV-19 Vaccination. *N*
1065 *Engl J Med* 384, 2124-2130. [10.1056/NEJMoa2104882](https://doi.org/10.1056/NEJMoa2104882).
- 1066 8. Lamprinou, M., Sachinidis, A., Stamoula, E., Vavilis, T., and Papazisis, G. (2023).
1067 COVID-19 vaccines adverse events: potential molecular mechanisms. *Immunol Res* 71,
1068 356-372. [10.1007/s12026-023-09357-5](https://doi.org/10.1007/s12026-023-09357-5).

- 1069 9. Buoninfante, A., Andeweg, A., Genov, G., and Cavaleri, M. (2024). Myocarditis
1070 associated with COVID-19 vaccination. *NPJ Vaccines* 9, 122. 10.1038/s41541-024-
1071 00893-1.
- 1072 10. Krumholz, H.M., Wu, Y., Sawano, M., Shah, R., Zhou, T., Arun, A.S., Khosla, P.,
1073 Kaleem, S., Vashist, A., Bhattacharjee, B., et al. (2023). Post-Vaccination Syndrome: A
1074 Descriptive Analysis of Reported Symptoms and Patient Experiences After Covid-19
1075 Immunization. *medRxiv*. 10.1101/2023.11.09.23298266.
- 1076 11. Semmler, A., Mundorf, A.K., Kuechler, A.S., Schulze-Bosse, K., Heidecke, H., Schulze-
1077 Forster, K., Schott, M., Uhrberg, M., Weinhold, S., Lackner, K.J., et al. (2023). Chronic
1078 Fatigue and Dysautonomia following COVID-19 Vaccination Is Distinguished from
1079 Normal Vaccination Response by Altered Blood Markers. *Vaccines (Basel)* 11.
1080 10.3390/vaccines11111642.
- 1081 12. Bruce K. Patterson, R.Y., Edgar B. Francisco, Emily Long, Amruta Pise, Eric Osgood,
1082 John Bream, Mark Kreimer, Devon Jeffers, Christopher Beaty, Richard Vander Heide,
1083 Jose Guevara-Coto, Rodrigo A Mora-Rodríguez Persistence of S1 Spike Protein in
1084 CD16+ Monocytes up to 245 Days in SARS-CoV-2 Negative Post COVID-19
1085 Vaccination Individuals with Post-Acute Sequelae of COVID-19 (PASC)-Like Symptoms.
- 1086 13. Trougakos, I.P., Terpos, E., Alexopoulos, H., Politou, M., Paraskevis, D., Scorilas, A.,
1087 Kastiritis, E., Andreakos, E., and Dimopoulos, M.A. (2022). Adverse effects of COVID-19
1088 mRNA vaccines: the spike hypothesis. *Trends Mol Med* 28, 542-554.
1089 10.1016/j.molmed.2022.04.007.
- 1090 14. Verbeke, R., Hogan, M.J., Lore, K., and Pardi, N. (2022). Innate immune mechanisms of
1091 mRNA vaccines. *Immunity* 55, 1993-2005. 10.1016/j.immuni.2022.10.014.
- 1092 15. Appledorn, D.M., Patial, S., McBride, A., Godbehere, S., Van Rooijen, N.,
1093 Parameswaran, N., and Amalfitano, A. (2008). Adenovirus vector-induced innate
1094 inflammatory mediators, MAPK signaling, as well as adaptive immune responses are
1095 dependent upon both TLR2 and TLR9 in vivo. *J Immunol* 181, 2134-2144.
1096 10.4049/jimmunol.181.3.2134.
- 1097 16. Ogata, A.F., Cheng, C.A., Desjardins, M., Senussi, Y., Sherman, A.C., Powell, M.,
1098 Novack, L., Von, S., Li, X., Baden, L.R., and Walt, D.R. (2022). Circulating Severe Acute
1099 Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Vaccine Antigen Detected in the
1100 Plasma of mRNA-1273 Vaccine Recipients. *Clin Infect Dis* 74, 715-718.
1101 10.1093/cid/ciab465.
- 1102 17. Cognetti, J.S., and Miller, B.L. (2021). Monitoring Serum Spike Protein with Disposable
1103 Photonic Biosensors Following SARS-CoV-2 Vaccination. *Sensors (Basel)* 21.
1104 10.3390/s21175857.
- 1105 18. Patterson, B.K., Francisco, E.B., Yogendra, R., Long, E., Pise, A., Rodrigues, H., Hall,
1106 E., Herrera, M., Parikh, P., Guevara-Coto, J., et al. (2021). Persistence of SARS CoV-2
1107 S1 Protein in CD16+ Monocytes in Post-Acute Sequelae of COVID-19 (PASC) up to 15
1108 Months Post-Infection. *Front Immunol* 12, 746021. 10.3389/fimmu.2021.746021.
- 1109 19. Assessment report_Spikevax. (2023).
1110 [https://www.ema.europa.eu/en/documents/assessment-report/spikevax-previously-](https://www.ema.europa.eu/en/documents/assessment-report/spikevax-previously-covid-19-vaccine-moderna-epar-public-assessment-report_en.pdf)
1111 [covid-19-vaccine-moderna-epar-public-assessment-report_en.pdf](https://www.ema.europa.eu/en/documents/assessment-report/spikevax-previously-covid-19-vaccine-moderna-epar-public-assessment-report_en.pdf).
- 1112 20. Theoharides, T.C. (2022). Could SARS-CoV-2 Spike Protein Be Responsible for Long-
1113 COVID Syndrome? *Mol Neurobiol* 59, 1850-1861. 10.1007/s12035-021-02696-0.
- 1114 21. Segal, Y., and Shoenfeld, Y. (2018). Vaccine-induced autoimmunity: the role of
1115 molecular mimicry and immune crossreaction. *Cell Mol Immunol* 15, 586-594.
1116 10.1038/cmi.2017.151.
- 1117 22. Klein, J., Wood, J., Jaycox, J.R., Dhodapkar, R.M., Lu, P., Gehlhausen, J.R.,
1118 Tabachnikova, A., Greene, K., Tabacof, L., Malik, A.A., et al. (2023). Distinguishing

- 1119 features of long COVID identified through immune profiling. *Nature* 623, 139-148.
1120 10.1038/s41586-023-06651-y.
- 1121 23. The Yale LISTEN Study. <https://medicine.yale.edu/ycci/listen-study/>.
- 1122 24. Ma, Q., Liu, J., Liu, Q., Kang, L., Liu, R., Jing, W., Wu, Y., and Liu, M. (2021). Global
1123 Percentage of Asymptomatic SARS-CoV-2 Infections Among the Tested Population and
1124 Individuals With Confirmed COVID-19 Diagnosis: A Systematic Review and Meta-
1125 analysis. *JAMA Netw Open* 4, e2137257. 10.1001/jamanetworkopen.2021.37257.
- 1126 25. El-Khoury, J.M., Schulz, W.L., and Durant, T.J.S. (2021). Longitudinal Assessment of
1127 SARS-CoV-2 Antinucleocapsid and Antispike-1-RBD Antibody Testing Following PCR-
1128 Detected SARS-CoV-2 Infection. *J Appl Lab Med* 6, 1005-1011. 10.1093/jalm/jfab030.
- 1129 26. Pirofski, L.A., and Casadevall, A. (2020). The state of latency in microbial pathogenesis.
1130 *J Clin Invest* 130, 4525-4531. 10.1172/JCI136221.
- 1131 27. Hodel, F., Naret, O., Bonnet, C., Brenner, N., Bender, N., Waterboer, T., Marques-Vidal,
1132 P., Vollenweider, P., and Fellay, J. (2022). The combined impact of persistent infections
1133 and human genetic variation on C-reactive protein levels. *BMC Med* 20, 416.
1134 10.1186/s12916-022-02607-7.
- 1135 28. Su, Y., Yuan, D., Chen, D.G., Ng, R.H., Wang, K., Choi, J., Li, S., Hong, S., Zhang, R.,
1136 Xie, J., et al. (2022). Multiple early factors anticipate post-acute COVID-19 sequelae.
1137 *Cell* 185, 881-895 e820. 10.1016/j.cell.2022.01.014.
- 1138 29. Rodriguez, L., Tan, Z., Lakshmikanth, T., Wang, J., Barcenilla, H., Swank, Z., Zuo, F.,
1139 Abolhassani, H., Pavlovitch-Bedzyk, A. J., Wang, C., Gonzalez, L., Mugabo, C. H.,
1140 Johnsson, A., Chen, Y., James, A., Mikes, J., Kleberg, L., Sundling, C., Björnson, M.,
1141 Nygren Bonnier, M., Ståhlberg, M., Runold, M., Björkander, S., Melén, E., Meyts, I., Van
1142 Weyenbergh, J., Hammarström, Q.-P., Davis, M. M., Walt, D. R., Landegren, N., COVID
1143 Human Genetic Effort, Aiuti, A., Casari, G., Casanova, J.-L., Jamouille, M., Bruchfeld, J.,
1144 Brodin, P. (2024). Restrained memory CD8+ T cell responses favors viral persistence
1145 and elevated IgG responses in patients with severe Long COVID.
- 1146 30. Didier, K., Bolko, L., Giusti, D., Toquet, S., Robbins, A., Antonicelli, F., and Servettaz, A.
1147 (2018). Autoantibodies Associated With Connective Tissue Diseases: What Meaning for
1148 Clinicians? *Front Immunol* 9, 541. 10.3389/fimmu.2018.00541.
- 1149 31. Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation
1150 network analysis. *BMC Bioinformatics* 9, 559. 10.1186/1471-2105-9-559.
- 1151 32. Jiang, R., Janssen, M.F.B., and Pickard, A.S. (2021). US population norms for the EQ-
1152 5D-5L and comparison of norms from face-to-face and online samples. *Qual Life Res*
1153 30, 803-816. 10.1007/s11136-020-02650-y.
- 1154 33. Mundorf, A.K., Semmler, A., Heidecke, H., Schott, M., Steffen, F., Bittner, S., Lackner,
1155 K.J., Schulze-Bosse, K., Pawlitzki, M., Meuth, S.G., et al. (2024). Clinical and Diagnostic
1156 Features of Post-Acute COVID-19 Vaccination Syndrome (PACVS). *Vaccines (Basel)*
1157 12. 10.3390/vaccines12070790.
- 1158 34. Scholkmann, F., and May, C.A. (2023). COVID-19, post-acute COVID-19 syndrome
1159 (PACS, "long COVID") and post-COVID-19 vaccination syndrome (PCVS, "post-
1160 COVIDvac-syndrome"): Similarities and differences. *Pathol Res Pract* 246. ARTN
1161 154497
1162 10.1016/j.prp.2023.154497.
- 1163 35. Kwan, A.C., Ebinger, J.E., Wei, J., Le, C.N., Oft, J.R., Zabner, R., Teodorescu, D.,
1164 Botting, P.G., Navarrette, J., Ouyang, D., et al. (2023). Author Correction: Apparent risks
1165 of postural orthostatic tachycardia syndrome diagnoses after COVID-19 vaccination and
1166 SARS-Cov-2 Infection. *Nat Cardiovasc Res* 2, 956. 10.1038/s44161-023-00339-2.
- 1167 36. Safavi, F., Gustafson, L., Walitt, B., Lehky, T., Dehbashi, S., Wiebold, A., Mina, Y., Shin,
1168 S., Pan, B., Polydefkis, M., et al. (2022). Neuropathic symptoms with SARS-CoV-2
1169 vaccination. *medRxiv*. 10.1101/2022.05.16.22274439.

- 1170 37. Davis, H.E., McCorkell, L., Vogel, J.M., and Topol, E.J. (2023). Long COVID: major
1171 findings, mechanisms and recommendations. *Nat Rev Microbiol* 21, 133-146.
1172 10.1038/s41579-022-00846-2.
- 1173 38. Yonker, L.M., Swank, Z., Bartsch, Y.C., Burns, M.D., Kane, A., Boribong, B.P., Davis,
1174 J.P., Loiselle, M., Novak, T., Senussi, Y., et al. (2023). Circulating Spike Protein
1175 Detected in Post-COVID-19 mRNA Vaccine Myocarditis. *Circulation* 147, 867-876.
1176 10.1161/CIRCULATIONAHA.122.061025.
- 1177 39. Proal, A.D., VanElzaker, M.B., Aleman, S., Bach, K., Boribong, B.P., Buggert, M.,
1178 Cherry, S., Chertow, D.S., Davies, H.E., Dupont, C.L., et al. (2023). SARS-CoV-2
1179 reservoir in post-acute sequelae of COVID-19 (PASC). *Nat Immunol* 24, 1616-1627.
1180 10.1038/s41590-023-01601-2.
- 1181 40. Ryu, J.K., Yan, Z., Montano, M., Sozmen, E.G., Dixit, K., Suryawanshi, R.K., Matsui, Y.,
1182 Helmy, E., Kaushal, P., Makanani, S.K., et al. (2024). Fibrin drives thromboinflammation
1183 and neuropathology in COVID-19. *Nature* 633, 905-913. 10.1038/s41586-024-07873-4.
- 1184 41. Grobbelaar, L.M., Venter, C., Vlok, M., Ngoepe, M., Laubscher, G.J., Lourens, P.J.,
1185 Steenkamp, J., Kell, D.B., and Pretorius, E. (2021). SARS-CoV-2 spike protein S1
1186 induces fibrin(ogen) resistant to fibrinolysis: implications for microclot formation in
1187 COVID-19. *Biosci Rep* 41. 10.1042/BSR20210611.
- 1188 42. Rodriguez, L., Tan, Z., Lakshmikanth, T., Wang, J., Barcenilla, H., Swank, Z., Zuo, F.,
1189 Abolhassani, H., Pavlovitch-Bedzyk, A.J., Wang, C., et al. (2024). Restrained memory
1190 CD8⁺ T cell responses favors viral persistence and elevated IgG responses
1191 in patients with severe Long COVID. medRxiv, 2024.2002.2011.24302636.
1192 10.1101/2024.02.11.24302636.
- 1193 43. Swank, Z., Senussi, Y., Manickas-Hill, Z., Yu, X.G., Li, J.Z., Alter, G., and Walt, D.R.
1194 (2023). Persistent Circulating Severe Acute Respiratory Syndrome Coronavirus 2 Spike
1195 Is Associated With Post-acute Coronavirus Disease 2019 Sequelae. *Clin Infect Dis* 76,
1196 e487-e490. 10.1093/cid/ciac722.
- 1197 44. Swank, Z., Borberg, E., Chen, Y., Senussi, Y., Chalise, S., Manickas-Hill, Z., Yu, X.G.,
1198 Li, J.Z., Alter, G., Henrich, T.J., et al. (2024). Measurement of circulating viral antigens
1199 post-SARS-CoV-2 infection in a multicohort study. *Clin Microbiol Infect* 30, 1599-1605.
1200 10.1016/j.cmi.2024.09.001.
- 1201 45. Peluso, M.J., Swank, Z.N., Goldberg, S.A., Lu, S., Dalhuisen, T., Borberg, E., Senussi,
1202 Y., Luna, M.A., Chang Song, C., Clark, A., et al. (2024). Plasma-based antigen
1203 persistence in the post-acute phase of COVID-19. *Lancet Infect Dis* 24, e345-e347.
1204 10.1016/S1473-3099(24)00211-1.
- 1205 46. McAlpine, L. (2024). Vascular Inflammation in Neuropsychiatric Long COVID. *Ann*
1206 *Neurol* 96, S261-S262.
- 1207 47. Duarte-Rey, C., Bogdanos, D.P., Leung, P.S.C., Anaya, J.M., and Gershwin, M.E.
1208 (2012). IgM predominance in autoimmune disease: Genetics and gender. *Autoimmun*
1209 *Rev* 11, A404-A412. 10.1016/j.autrev.2011.12.001.
- 1210 48. Wang, H., Zang, C., Ren, M., Shang, M., Wang, Z., Peng, X., Zhang, Q., Wen, X., Xi, Z.,
1211 and Zhou, C. (2020). Cellular uptake of extracellular nucleosomes induces innate
1212 immune responses by binding and activating cGMP-AMP synthase (cGAS). *Sci Rep* 10,
1213 15385. 10.1038/s41598-020-72393-w.
- 1214 49. Uzawa, A., Oertel, F.C., Mori, M., Paul, F., and Kuwabara, S. (2024). NMOSD and
1215 MOGAD: an evolving disease spectrum. *Nat Rev Neurol* 20, 602-619. 10.1038/s41582-
1216 024-01014-1.
- 1217 50. Flow Repository-FR-FCM-Z8FZ
- 1218 51. Hugo Health. <https://hugo.health/>.

- 1219 52. World Medical, A. (2013). World Medical Association Declaration of Helsinki: ethical
1220 principles for medical research involving human subjects. *JAMA* 310, 2191-2194.
1221 10.1001/jama.2013.281053.
- 1222 53. von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gotsche, P.C., Vandenbroucke,
1223 J.P., and Initiative, S. (2007). The Strengthening the Reporting of Observational Studies
1224 in Epidemiology (STROBE) statement: guidelines for reporting observational studies.
1225 *Ann Intern Med* 147, 573-577. 10.7326/0003-4819-147-8-200710160-00010.
- 1226 54. Scientific Image and Illustration Software-Biorender. <https://www.biorender.com/>.
- 1227 55. Filardi, B.A., Monteiro, V.S., Schwartzmann, P.V., do Prado Martins, V., Zucca, L.E.R.,
1228 Baiocchi, G.C., Malik, A.A., Silva, J., Hahn, A.M., Chen, N.F.G., et al. (2023). Age-
1229 dependent impairment in antibody responses elicited by a homologous CoronaVac
1230 booster dose. *Sci Transl Med* 15, eade6023. 10.1126/scitranslmed.ade6023.
- 1231 56. Kamath, K., Reifert, J., Johnston, T., Gable, C., Pantazes, R.J., Rivera, H.N., McAuliffe,
1232 I., Handali, S., and Daugherty, P.S. (2020). Antibody epitope repertoire analysis enables
1233 rapid antigen discovery and multiplex serology. *Sci Rep* 10, 5294. 10.1038/s41598-020-
1234 62256-9.
- 1235 57. Haynes, W.A., Kamath, K., Waitz, R., Daugherty, P.S., and Shon, J.C. (2021). Protein-
1236 Based Immunome Wide Association Studies (PIWAS) for the Discovery of Significant
1237 Disease-Associated Antigens. *Front Immunol* 12, 625311. 10.3389/fimmu.2021.625311.
- 1238 58. Mu-Mosley, H., von Itzstein, M.S., Fattah, F., Liu, J., Zhu, C., Xie, Y., Wakeland, E.K.,
1239 Park, J.Y., Kahl, B.S., Diefenbach, C.S., and Gerber, D.E. (2024). Distinct autoantibody
1240 profiles across checkpoint inhibitor types and toxicities. *Oncoimmunology* 13, 2351255.
1241 10.1080/2162402X.2024.2351255.
- 1242 59. Ghosh, N., Postow, M., Zhu, C., Jannat-Khah, D., Li, Q.Z., Vitone, G., Chan, K.K., and
1243 Bass, A.R. (2022). Lower baseline autoantibody levels are associated with immune-
1244 related adverse events from immune checkpoint inhibition. *J Immunother Cancer* 10.
1245 10.1136/jitc-2021-004008.
- 1246 60. Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *J Roy Stat Soc*
1247 B 58, 267-288. DOI 10.1111/j.2517-6161.1996.tb02080.x.
- 1248