

1 **COHORT PROFILE: IMMUNE RESPONSES TO SARS-COV-2 VACCINATION AND INFECTION IN**  
2 **A LONGITUDINAL SAMPLING AMIDST THE COVID-19 PANDEMIC (LONGTONG-SARS2) IN**  
3 **MALAYSIA**

4

5 Naim Che-Kamaruddin<sup>1\*</sup>, Jefree Johari<sup>1</sup>, Hasmawati Yahaya<sup>1</sup>, Huy C. Nguyen<sup>2</sup>, Andrew G.  
6 Letizia<sup>2</sup>, Robert D. Hontz<sup>2</sup>, Sazaly AbuBakar<sup>1\*</sup>

7

8 <sup>1</sup>*Tropical Infectious Diseases Research and Education Centre (TIDREC), Higher Institution*  
9 *Centre of Excellence (HICoE), Universiti Malaya, 50603 Kuala Lumpur, Malaysia*

10 <sup>2</sup>*U.S. Naval Medical Research Unit INDO PACIFIC, Singapore*

11

12 **Email addresses:**

13 Naim Che-Kamaruddin (naimck@um.edu.my)

14 Jefree Johari (jefree@um.edu.my)

15 Hamawati Yahaya (hasmy@um.edu.my)

16 Huy C. Nguyen (huy.c.nguyen3.mil@health.mil)

17 Andrew G. Letizia (andrew.g.letizia.mil@health.mil)

18 Robert D. Hontz (hontzRD@state.gov)

19 Sazaly AbuBakar (sazaly@um.edu.my)

20 *\*corresponding authors*

21

22 **Word count:** 5,900

23

24

25

26

27

28 **ABSTRACT**

29 **Purpose:** This prospective, longitudinal study aims to evaluate the durability and  
30 functionality of SARS-CoV-2 Ancestral strain (Wuhan-Hu-1)-specific immune responses  
31 induced by COVID-19 vaccination and natural infection over a 12-month period. This article  
32 reviews the study protocol, design, methodology, ongoing data collection, analysis  
33 procedures, and demographic characteristics of the cohort enrolled.

34

35 **Participants:** Between March 2021 and May 2022, 400 participants were enrolled with a 12-  
36 month follow-up, concluding in May 2023. Two main groups of participants: (1) serologically  
37 SARS-CoV-2-naïve individuals receiving the BNT162b2 primary series vaccination (referred  
38 to as VAC) and (2) those who recently recovered from COVID-19 infection within 30 days,  
39 regardless of vaccination history (referred to as COV). Additionally, a subset of 45  
40 participants with selected COVID-19 exposure histories provided peripheral blood  
41 mononuclear cells (PBMCs) for cross-sectional analysis six months after enrollment.

42

43 **Findings to date:** Out of 400 participants, 66.8% (n=267) completed the follow-up. Among  
44 them, 52.8% (n=141) were in VAC, and 47.2% (n=126) were in COV. As the study progressed,  
45 we acknowledged cross-over between initial groups, leading to restructuring into five  
46 revised groups based on sequential exposure events. Sociodemographic factors revealed  
47 statistically significant age distribution differences ( $p=0.001$ ) in both initial and revised  
48 groups, with no significant differences observed for sex.

49

50 **Future plans:** LONGTONG-SARS2 assesses the host-pathogen interactions central to the  
51 development of COVID-19 immunity. With enrollment spanning two years of the pandemic,  
52 most participants exhibited mixed SARS-CoV-2 exposures—via vaccination and infection—  
53 resulting in diverse subgroups of interest. Notably, the inclusion of SARS-CoV-2-naïve, pre-  
54 exposure serum samples allowed for robust comparator and reduced potential biases.  
55 Ongoing analyses will include serology kinetics, memory cells ELISpots, B cells repertoire  
56 analysis, cytokine/chemokine profiling, and proteomic pathway to comprehensively  
57 examine the immune response against the SARS-CoV-2, thus informing and potentially  
58 predicting dynamic longitudinal responses against new more transmissible, immune-evasive  
59 SARS-CoV-2 variants.

60

61 **Keywords:** cellular immunity, correlate of protection, COVID-19, humoral response, serology

62

### 63 **STRENGTH AND LIMITATIONS OF THIS STUDY**

64

65 - LONGTONG-SARS2 is a prospective longitudinal study that comprehensively evaluates the  
66 SARS-CoV-2 immune response among a diverse group of individuals, stratified based on the  
67 sequential order of SARS-CoV-2 exposure events, whether from COVID-19 vaccination or  
68 infection.

69

70 - Pre-vaccination serum samples were collected from serologically SARS-CoV-2 naive  
71 individuals scheduled to receive the BNT162b2 primary series vaccination during the initial  
72 mass COVID-19 vaccination phase in Malaysia in early 2021.

73

74 - The longitudinal serum sample collection spanned two years of the COVID-19 pandemic,  
75 from March 2021 to May 2023. This extended duration allows for robust monitoring of the  
76 immune response against SARS-CoV-2 variants in comparison to the ancestral strain.

77

78 - There is a risk of misclassification of some individuals' SARS-CoV-2 exposure status through  
79 serology, as certain sampling timepoints had intervals of three months. Additionally, our  
80 study relies on self-reported data through the *MySejahtera* application (Malaysia's  
81 electronic medical record by the Ministry of Health) for second confirmation, potentially  
82 leading to underdiagnosed and underreported cases of asymptomatic infection.

83

### 84 **1.0 INTRODUCTION**

85

86 Coronavirus disease 2019 (COVID-19) is an ongoing viral respiratory illness caused by Severe  
87 Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). SARS-CoV-2 is a positive-sense,  
88 single-stranded RNA virus. This Risk Group 3 (RG3) virus is phylogenetically related to other  
89 Betacoronaviruses which can cause respiratory illnesses in humans, including Severe Acute  
90 Respiratory Syndrome Coronavirus 1 (SARS-CoV-1) and the Middle East Respiratory  
91 Syndrome Coronavirus (MERS-CoV) [1]. As of 31 December 2023, SARS-CoV-2 has been

92 responsible for more than 773 million COVID-19 cases and more than 7 million deaths  
93 globally, with the actual death toll likely to be higher, largely due to the lack of reliable  
94 mortality surveillance and tracking systems in many countries and instances where  
95 individuals succumbed to COVID-19 before being tested for the virus [2]. Most patients  
96 infected with SARS-CoV-2 experienced asymptomatic, mild, or moderate symptoms, and  
97 recovered without hospitalization, while a minority developed severe or critical clinical  
98 manifestations [3].

99

100 The effectiveness of SARS-CoV-2 vaccines has been well documented [4–6]. However,  
101 waning antibody titers which has been documented following three to four months post-  
102 infection and reports of reduced vaccine efficacy against emerging SARS-CoV-2 variants  
103 underscore the need for ongoing assessment of continuous booster immunization [7].  
104 Previous studies suggested that the risk of reinfection is attributed to the waning of  
105 neutralizing antibodies (NAb) [4–7]. According to these earlier studies, antibodies are only  
106 detected in 60–85% of convalescent individuals, despite detectable levels in 80–95% within  
107 two weeks post-symptom onset [8–10]. Considering the serologic kinetics, booster  
108 administration increased the serum NAb titer and cross-reactivity against emerging Omicron  
109 variants [11]. Although booster immunization has been shown to stimulate antigen-specific  
110 antibody production, additional more comprehensive studies are needed to better  
111 understand the multifaceted immunological mechanisms that correlate with clinical  
112 protection following vaccination and/or natural infection, particularly in the context of the  
113 continuing emergence of new immune-invasive SARS-CoV-2 variants [12].

114

115 SARS-CoV-2 exposure from vaccination and/or infection elicits an adaptive immune  
116 response, engaging both humoral (primarily antibody-mediated) and cellular immunity,  
117 which includes the activation of B and T cells [13]. This dual response establishes long-term  
118 memory within the immune system. Upon re-exposure to SARS-CoV-2, the memory B and T  
119 cells expand to mount a rapid and targeted defense [6]. Immune memory cells,  
120 demonstrated to be more durable, contribute to long-term immunity against variants, even  
121 as antibody titers contract [5,6,14]. Evidence of immune imprinting, wherein initial  
122 exposures effectively prime B cells but potentially limit the development of novel memory B  
123 cells and NAb against emerging SARS-CoV-2 variants, poses a significant challenge to

124 maintaining vaccine efficacy amid the ongoing emergence of novel SARS-CoV-2 variants  
125 [15,16]. While the mechanisms of how the immune system recognizes and responds to  
126 SARS-CoV-2 variants are still being elucidated, both humoral and cellular immunities play  
127 important roles in conferring protection [17]. Therefore, it is imperative to investigate the  
128 impact of varied SARS-CoV-2 exposure histories on shaping the immune responses to SARS-  
129 CoV-2 [16], especially on a regional basis. This is attributed to the variation in immune  
130 response across population, with factors such as genetic, environmental, and lifestyle being  
131 unique to each region [18]. Moreover, the prevalence of SARS-CoV-2 variants and  
132 vaccination strategies in each region can significantly impact viral transmission dynamics.  
133 This information becomes crucial for informing public health policies, particularly in the  
134 context of emerging highly transmissible variants [19].

135

136 The present study seeks to establish a comprehensive longitudinal prospective cohort  
137 comprised of participants exposed to SARS-CoV-2 vaccination and/or infection to define the  
138 epidemiological and immunological effects of SARS-CoV-2 in Malaysia. To our knowledge,  
139 Malaysia lacks a cohort with a similar extensive design, involving longitudinal sampling up to  
140 12 months, spanning from the early phase of the pandemic to the transitioning to an  
141 endemic phase. This study fills a crucial gap by providing insights into COVID-19 humoral  
142 and cellular immune responses specific to Malaysia. Through investigation of antibody  
143 kinetics, T and B cell studies, and adaptive immune responses, our initiative not only  
144 contributes valuable data for informing public health strategies but also has the potential to  
145 serve as a reference for future research in the region. This article describes the protocol and  
146 cohort profile of **LONGitudinal sTudy Of immuNe response aGainst SARS-CoV-2**  
147 (LONGTONG-SARS2).

148

## 149 **2.0 STUDY OBJECTIVES**

150

151 The overall aim of “**LONGitudinal sTudy Of immuNe response aGainst SARS-CoV-2**”  
152 (LONGTONG-SARS2) is to evaluate the durability of SARS-CoV-2 specific antibodies, and the  
153 functionality of the specific memory B and T cells from a prospective, longitudinal sampling  
154 for up to 12 months. The specific objectives include:

155

- 156 1. To analyze the kinetics of anti-SARS-CoV-2 antibodies (IgM, IgA, IgG, NAb) response  
157 for up to 12 months by ELISA assays for each group.
- 158 2. To assess the functional immune memory response of SARS-CoV-2-specific memory  
159 B and T cells by ELISpot assays.
- 160 3. To characterize the interaction of cytokines and immune cells through bead-based  
161 analytes.
- 162 4. To profile the serum protein pathways before and after SARS-CoV-2 vaccination or  
163 infection from bottom-up proteomic analysis.
- 164 5. To identify the distribution and incidence of SARS-CoV-2 exposures by duration since  
165 the primary series of vaccination doses, frequency of infection(s), and date of  
166 exposure in the study population by descriptive analysis.
- 167 6. To evaluate the association of antibody distribution and baseline demographic,  
168 comorbidities, as well as symptoms/adverse effects following immunization (AEFI)  
169 observed during enrollment, subsequent to SARS-CoV-2 exposures via vaccination or  
170 infection by generalized estimating equation (GEE) analysis.

171

### 172 **3.0 COHORT DESCRIPTION**

173

#### 174 **3.1 Terminology definitions:**

175 The terminology used in the present study are:

- 176 • "Vaccination" - Participants were classified as vaccinated if they received COVID-19  
177 vaccination regardless of number of doses. Proof of vaccination is a mandatory  
178 requirement for inclusion in any vaccination data collection within the scope of the  
179 present study. Participants who received the primary vaccination series were  
180 categorized as fully vaccinated two weeks after the last dose, while those receiving  
181 additional vaccine doses after the primary series were classified as boosted.
- 182 • "Infection" - Participants were classified as infected if they presented proof of a  
183 positive case through PCR or rapid test results in either their records (as well as from  
184 the electronic medical record by Ministry of Health, *MySejahtera* application) OR  
185 demonstrated at least a 4-fold increase in IgG or IgM levels from the consecutive  
186 sampling timepoints. Individuals exhibiting a 4-fold increase in IgG or IgM without

187 documented proof (due to the absence of symptoms) were considered to have  
188 asymptomatic infection.

- 189 • "Reinfection" - Reinfection status for participants was documented if multiple  
190 occurrences of COVID-19 infections were recorded in the electronic medical record  
191 OR if there was at least a four-fold increase in IgG or IgM levels from the consecutive  
192 sampling timepoints even without the records.
- 193 • "Breakthrough infection (BTI)" - The BTI status for participants was documented if  
194 the date of infection occurred after Day-15 of the second dose of COVID-19  
195 vaccination, indicating completion of the vaccination regimen.
- 196 • "Exposure" - Vaccination, infection, reinfection, and BTI were considered as  
197 exposure in the present study.
- 198 • "Cross-exposure" - Participants were defined to have cross-exposure if they  
199 experienced both COVID-19 vaccination and infection, irrespective of the sequence  
200 of exposure.

201

### 202 **3.2 Study design**

203

204 The study design involved a longitudinal study with enrollment of two groups of different  
205 SARS-CoV-2 exposures, which are (1) the serologically SARS-CoV-2-naïve individuals who  
206 were scheduled to receive primary series vaccination with BNT162b2 (VAC) and (2) post-  
207 COVID-19 infected group (COV) regardless of their vaccination histories. The participants  
208 were informed beforehand about all aspects of the study, and individuals who fulfilled the  
209 inclusion criteria were given an informed consent form and a case report form (CRF)  
210 adhering ethical approval (MREC-UMMC SID: 2021226-9886).

211

212 The inclusion criteria are as follows:

213

- 214 ☐ Seven years or older, AND
- 215 ☐ Healthy individuals without active SARS-CoV-2 infection at the time of enrollment,  
216 AND

- 217       ☐ Healthy individuals without immunosuppressive conditions such as HIV and cancer  
218       patients, AND EITHER  
219       ☐ Scheduled for the primary series vaccination with BNT162b2 with a negative baseline  
220       SARS-CoV-2 anti-S1 IgG serology, corresponding to no previous SARS-CoV-2 exposure  
221       (to be enrolled in the VAC group), OR  
222       ☐ Had been infected with SARS-CoV-2 within 30 days and able to provide proof of  
223       infection either from rapid test kit or PCR (to be enrolled in the COV group),  
224       ☐ AND agreed to give voluntary informed consent to participate in the study  
225

226 The exclusion criteria are as follows:

- 227       ☐ Individuals with immunosuppressive conditions such as HIV and cancer patients, OR  
228       ☐ Individuals unable to attend the scheduled follow-up appointments  
229

230 Epidemiological data obtained in the CRF include:

- 231       ☐ Demographics - *Age, gender, body mass index (BMI), medical comorbidities*  
232       ☐ Exposure history - *Date of vaccinations, COVID-19 vaccine types, date of COVID-19*  
233       *infections*  
234       ☐ Clinical manifestations experienced - *Adverse effect following immunization (AEFI)*  
235       *and symptoms following COVID-19 infection*  
236

### 237 **3.3 Sample size estimation**

238

239 A sample size of 400 participants (N=400) was targeted to achieve statistical power as  
240 calculated using GLIMMPSE software (<https://glimmpse.samplesizeshop.org/>), taking into  
241 consideration the repeated measures and longitudinal design [9,10,20].  
242

243 In GLIMMPSE, the Hotelling-Lawley Trace test was utilized with a conservative Type 1 error  
244 rate of 0.5, with the total sample size for the analysis being 66 participants. An estimated  
245 30% attrition rate was used, and an adjusted estimation formula of  $66 / (1-0.3)$  was applied  
246 to accommodate this attrition, resulting in a minimum total sample size requirement of 95  
247 participants. For the present study, the total of 400 participants were chosen to increase  
248 the statistical power and accommodate potential loss of follow-up participation.



249

### 250 **3.4 Enrollment**

251

252 Following all required ethics approvals, eligible participants were enrolled. Enrollment  
253 material (i.e., advertisements, posters, and flyers) were posted on social and printed media.  
254 The target study participants primarily comprised of individuals from the public residing in  
255 the greater Kuala Lumpur area (Supplemental Figure 1). Due to the restrictions imposed by  
256 the nationwide MCO, we utilized both convenience sampling, selecting participants based  
257 on their accessibility and proximity to the study site, and snowball sampling, where existing  
258 participants referred others for the study. Participant enrollment commenced in March  
259 2021, coinciding with the initial phase of COVID-19 vaccine distribution in Malaysia. Initially,  
260 potential participants were provided with comprehensive information on the study, which  
261 included the study purpose, blood collection procedures, and the prerequisite follow-ups.  
262 Consented participants completed informed consent. Visit schedules for sample collection  
263 were sent via an electronic appointment card. Collected blood samples were transported on  
264 ice to the laboratory, where the serum samples were processed and stored in -80°C freezer  
265 until needed for laboratory analysis.

266

267 Enrollment closed at the end of July 2022, and the blood sampling ended at the end of May  
268 2023.

269

### 270 **3.5 Blood sample collection**

271

272 Longitudinal blood sampling for up to 12 months was required for all participants following  
273 the timepoints described in Figure 1. Participants in the VAC group were scheduled for a  
274 total of 9 sampling timepoints, spanning from pre-vaccination or baseline to Day-360 days  
275 post the first vaccine dose. Meanwhile, the COV group had 7 scheduled timepoints, initiated  
276 from the first month of infection. Notably, sample collection timepoints for both groups  
277 coincided on Day-14 post first exposure, Day-90, Day-180, Day-270, and Day-360 post-  
278 vaccination/infection. A sampling window of  $\pm 7$  days was applied to the designated  
279 sampling timepoints for participants' flexibility and availability. One (5 mL) tube with a  
280 separator gel and clot activator for whole blood samples was drawn for each timepoint.

281 Blood samples were processed with serum extracted from each sample for the serology  
282 testing.

283

284 A cross-sectional collection of Peripheral Blood Mononuclear Cells (PBMCs) was conducted  
285 from Day-180 to Day-360 to assess the cellular immune response. Among the 400  
286 participants, 45 were randomly selected and consented for PBMCs collection. To establish a  
287 control group for cellular immune response analysis, an additional 5 participants who had  
288 no COVID-19 vaccination documented in a local electronic record, *MySejahtera* and were  
289 not previously involved in the LONGTONG-SARS2 serology study, were included. Details of  
290 the vaccination types and the frequency of SARS-CoV-2 infection were collected to form  
291 groups based on these parameters (details are shown in Supplemental Table 3). For the  
292 PBMCs collection, four (8 mL) lymphocyte isolation tubes (BD Vacutainer® CPT™  
293 Mononuclear Cell Preparation Tube) were used. The PBMCs were isolated and assessed for  
294 cell count estimation and viability using BioRad TC20™ Cell Counter.

295

### 296 **3.6 Participant and public involvement**

297

298 Participants or the public were not involved in the design, or conduct, or reporting, or  
299 dissemination plans of our study.

300

## 301 **4.0 LABORATORY INVESTIGATIONS**

302

303 There are five major components of laboratory analysis: i) serology monitoring, ii) B- and T-  
304 memory cell functionality, iii) B lymphocyte population counting, iv) cytokine profiling, and  
305 v) proteomic profiling (Figure 2). Briefly, the serology monitoring involves the longitudinal  
306 quantification for specific antibodies against SARS-CoV-2. The assessment of B- and T-  
307 memory cells functionality seeks to understand how well the immune memory responds  
308 upon re-exposure to SARS-CoV-2 antigens. B lymphocyte population counting intends to  
309 elucidate the heterogeneity in B cells population following the re-exposure, while cytokine  
310 profiling involves identifying key signaling molecules that are upregulated or downregulated  
311 as part of the immune response. Additionally, proteomic profiling provides a broader view  
312 of protein expression patterns during the immune response. Profiling the various arms of

313 the immune system will help establish a comprehensive analysis for a holistic understanding  
314 of how the immune system responds against SARS-CoV-2.

315

#### 316 **4.1 Serology monitoring**

317

318 Serological testing of serum samples employs four different commercially available anti-  
319 SARS-CoV-2 ELISA kits targeting either receptor-binding-domain (RBD) or spike (S) protein.

320 The kits used in the present study include detection of IgM (WANTAI, Beijing Wantai

321 Biological Pharmacy Ent.), IgA, IgG, and NAb (EUROIMMUN, Lubeck, Germany). Serologic

322 testing was conducted at every time point, following the manufacturer's protocols. The

323 threshold for positivity of RBD-specific IgM is defined at an optical density of 0.105 OD at a

324 wavelength of 450 nm ( $OD_{450}$ ), with a reference at 650 nm ( $OD_{650}$ ). For S1-specific IgA,

325 semiquantitative results are derived by calculating the OD ratio using the samples and a

326 provided standard. A positivity cut-off value of 1.1 OD ratio is established to interpret

327 results. For S1-specific IgG, quantitative determination of binding antibody units (BAU/mL) is

328 obtained using the standards provided by the manufacturer with a positivity cut-off value of

329 35.2 BAU/mL. Serum NAb activity against S1/RBD-Wuhan-Hu-1 SARS-CoV-2 (Wuhan

330 wildtype) is measured in a surrogate Virus neutralization Test (sVNT) based on competitive

331 ELISA. According to the manufacturer sheets, the sensitivity of these commercial ELISA kits

332 ranges from 86.1% to 96.9%, and the specificity ranges from 98.3% to 99.8%.

333

#### 334 **4.2 Memory cells functionality through ELISpot**

335

336 The functionality of memory B and T cells is analyzed using PBMCs stimulated against SARS-

337 CoV-2 antigens in vitro. Commercial ELISpot assays (Mabtech, Sweden) quantify the

338 production of RBD-specific IgG in memory B cells. The assays assess the production of

339 cytokines, particularly interferon-gamma ( $IFN-\gamma$ ) in memory T cells. Briefly,  $2 \times 10^5$  cells are

340 incubated with SARS-CoV-2 peptides for 24 to 72 hours, and spot development is performed

341 according to the manufacturer's protocol. Unstimulated controls for each sample function

342 as the internal reference, enabling a comparative analysis of stimulated changes using the

343 individual's baseline as a reference point. Spots are counted using an ImmunoSpot® analyzer

344 and converted into spots forming unit per million cells (SFU/mL) for standardization. The

345 collected data is analyzed to determine the frequency of cytokine-producing memory T cells  
346 and IgG-secreting memory B cells, as previously described [6].

347

### 348 **4.3 B lymphocytes subsets response through flow-cytometric analyses**

349

350 PBMCs were analyzed using flow cytometry to discern the frequency of distinct B cell  
351 subpopulations and their potential correlation with cell functionality. Flow cytometric  
352 analyses were designed in the present study to characterize the B cell response include cell  
353 viability (7-AAD), mature B cell lineage (CD19+ CD20+), transitional B cell (CD38 dim/-  
354 CD24+), antibody-secreting cell (CD38+ CD24- CD27+ CD20-), CD27 and IgD for atypical  
355 memory B cells, memory B cells, innate-like memory B cells, naïve B cells. Spike and RBD-  
356 specific SARS-CoV-2 markers, including BB515 Streptavidin BV421 Streptavidin for double  
357 spike+, BB515 Streptavidin BV421 Streptavidin for double spike+ with spike RBD+, and IgM  
358 with IgG to distinguish the primary and secondary response in the stimulated samples.  
359 Previous literature on cellular immunity response against SARS-CoV-2 were referred for the  
360 cells subset as previously describe [5,21]. A BD FASCLytic™ cell analyzer was be used to  
361 characterize the samples.

362

### 363 **4.4 Cytokine profiling**

364

365 In the subset of 45 selected participants, including the 5 unvaccinated control participants,  
366 cytokine levels from the supernatant of cells suspension are assessed using the Bio-Plex  
367 multiplex magnetic bead-based human cytokine assay (Bio-Rad, CA, USA) with the 27-Plex  
368 Screening Panel (#M500KCAFOY). The Bio-Plex 200 Reader measures median fluorescence  
369 intensities, and each sample undergoes triplicate analysis. Data acquisition and analysis are  
370 performed using Bio-Plex Manager Software. Standard curves for each cytokine are  
371 established using the provided manufacturer's standards according to manufacturer's  
372 instruction and as described previously [22].

373

374 In this study, we focused on 27 unique cytokines which involves in inflammatory responses  
375 and initiate cellular immune signals: eotaxin, fibroblast growth factor (FGF), granulocyte  
376 colony-stimulating factor (G-CSF), interferon- $\gamma$  (IFN- $\gamma$ ), IL-1 $\beta$ , interleukin 1 receptor

377 antagonist (IL-1RA), IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12(p70), IL-13, IL-15, IL-17A,  
378 interferon- $\gamma$ -inducible protein (IP-10), monocyte chemoattractant protein-1 (MCP-1, MCAF),  
379 macrophage inflammatory protein-1 $\alpha$  (MIP-1 $\alpha$ ), MIP-1 $\beta$ , platelet-derived growth factor-BB,  
380 regulated upon activation, normal T-cell expressed and presumably secreted (RANTES),  
381 tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ), and vascular endothelial growth factor-A (VEGF).

382

#### 383 **4.5 Proteome Profiling**

384

385 Longitudinal serum samples were used to study the host proteomic response against COVID-  
386 19 vaccination and/or infection. During the protein preparation, high abundance proteins  
387 were depleted by a commercial kit, High Select™ Depletion Spin Columns (Thermo Scientific,  
388 USA), and protein digestion was performed using EasyPep™ MS Sample Prep Kits (Thermo  
389 Scientific, USA). Protein identification and quantification are performed using the  
390 untargeted Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) Q-TOF (Agilent  
391 Technologies 6520, USA) system. Samples were identified using data-independent  
392 acquisition mode. Spectra from each fraction are explored in the protein database, and the  
393 results are imported to generate a library to identify differentially expressed proteins (DEPs)  
394 that significantly vary between samples. The functional significance of DEPs is explored  
395 using bioinformatics tools, including Gene Ontology (GO) and KEGG pathway analysis  
396 following protocols of available study [23].

397

#### 398 **4.6 Data analysis and modeling plan**

399

400 At the time of study conceptualization during the early phase of the pandemic, the initial  
401 objective was to compare the immune response between participants receiving only the  
402 BNT162b2 vaccine with no infection to those post-COVID-19 infection without vaccination,  
403 focusing on (1) measurement of serum-neutralizing antibodies, IgG, IgA, IgM response  
404 against SARS-CoV-2 for up to 12 months and (2) evaluation of the long term-memory B and  
405 T lymphocyte-mediated immune response. The inclusion of only BNT16b2 vaccine in the  
406 present study was informed by its predominant primary series vaccine used in Malaysia.  
407 Throughout the pandemic, the study analysis plans had to undergo continuous adjustments  
408 and adaptation to account for group overlap due to varying accumulation of SARS-CoV-2

409 exposures outside of the control of the study investigators, such as the implementation of  
410 the government's vaccine rollout plan and the emergence of new more transmissible and  
411 immune-evasive variants that cause breakthrough infections. To accommodate these  
412 limitations, additional data on COVID-19 vaccination, booster doses, and COVID-19 infection  
413 histories were requested from the participants at the end of their last sampling timepoints.  
414 These data were collected from a local government-sponsored and mandated electronic  
415 record, the *MySejahtera* application available on all participants' mobile devices to limit  
416 recall bias. The additional data were utilized to determine cross-exposure in participants,  
417 specifically examining the immune response against SARS-CoV-2 induced by either  
418 vaccination alone, infection alone, or a combination of both. Only participants who provided  
419 proof of vaccination certificate and diagnostic results from rapid test kits or PCR were  
420 included in the final analysis. Additionally, 4-fold increment increases in anti-SARS-CoV-2 S1  
421 IgG over consecutive sampling timepoints (outside of the vaccination period) were  
422 monitored to account for asymptomatic infections. Participants in the VAC group with  
423 COVID-19 infection are grouped based on breakthrough infection (BTI) status. Plans to  
424 analyze these results are in progress and will be reported in future publications.

425

426 A complete-case analysis approach was chosen, wherein only participants who completed  
427 the entire follow-up period for blood sampling and provided fully completed Case Report  
428 Forms (CRFs) were included in the analysis. The assumption was that the excluded  
429 participants were represented randomly, indicating that the missing data was categorized as  
430 Missing Completely at Random (MCAR) data [24] and therefore the excluded data will not  
431 affect the statistical analyses.

432

433 Descriptive statistics were employed to summarize the characteristics of study participants  
434 and relevant variables, providing an overview of data distribution and cohort demographics.  
435 Age groups were categorized according to common life stages, which are group  $\leq 30$  years  
436 old (teenagers and young adults), 31-40 years old (adults),  $> 40$  years old (late adults to  
437 seniors/elderly), respectively. Logistic regression will be used to assess the association  
438 between categorical outcomes (e.g., symptoms/AEFI severity, antibody waning status) and  
439 predictor variables (e.g., vaccination types, frequency of infection), while generalized  
440 estimating equation (GEE) models will be used to account for longitudinal data with

441 repeated measurements, enabling the examination of changes in immune response markers  
442 and BTIs/reinfections over time. Correlation tests explore relationships between continuous  
443 variables, revealing the strength and direction of factor associations. Additionally, survival  
444 analysis is employed for time-to-breakthrough infection/reinfection, and cluster analysis  
445 identifies distinct participant subgroups based on immune response profiles or other  
446 characteristics such as the SARS-CoV-2 vaccination types and infection frequency. The  
447 analyses are performed using established statistical software (e.g., R, SPSS, GraphPad  
448 Prism).

449

## 450 **5.0 FINDINGS TO DATE**

451

452 Initially 184 SARS-CoV-2 naïve, SARS-CoV-2 anti-S1 IgG seronegative participants were  
453 enrolled into the VAC group in March 2021 and April 2021 before their scheduled BNT162b2  
454 vaccination (Supplemental Figure 2). This was made possible by our university hospital's  
455 designation as a vaccination distribution center, allowing us to enroll study participants even  
456 during MCOs. Starting in May 2021, however, there was a decrease in the number of SARS-  
457 CoV-2 naïve participants due to the Malaysian government's mass vaccination campaign.  
458 This decline presented challenges in meeting our initial target of 200 participants for the  
459 VAC group.

460

461 In contrast, the enrollment for the COV group encountered a slowdown in the initial phase  
462 of the study, leading to an extension. This delay was prompted by unexpected restrictions  
463 imposed during the MCO implemented by the Malaysian government between March 2020  
464 and December 2021 including limitations on gathering, mobility, and international travel,  
465 along with the closure of government, business, and educational institutions to prevent the  
466 spread of SARS-CoV-2. However, enrollment for the COV group experienced a significant  
467 increase in March 2022, coinciding with the emergence of the Omicron variant. Despite the  
468 challenges, we successfully enrolled 216 COV group participants with post-COVID-19 natural  
469 infection. An additional 16 participants were included to compensate for the participation  
470 numbers in the VAC group.

471

472 The decrease in retention rate started at Day-180 of sample collection (Supplemental Table  
473 1). Through our observations and interactions with participants, the primary reason for  
474 study withdrawal was the resumption of their usual daily responsibilities during the  
475 transition from the pandemic to the endemic phase. Consequently, shifting priorities among  
476 participants led to a diminished interest in continuing their participation in the study. By the  
477 final follow-up on Day-360, the total retention rate throughout all sampling timepoints was  
478 70% (n=280/400) which providing a sample at every single sampling timepoint.

479

### 480 **5.1 Enrollment strategies**

481

482 Given the ongoing COVID-19 pandemic during the study, we informed the participants of  
483 their serology test results to make them aware of their status against SARS-CoV-2. By  
484 sharing serology results, participants were likely motivated to remain engaged in the study,  
485 as they recognized the importance of continuous surveillance.

486

487 In addition, we also implemented a mobile follow-up service to enhance participant  
488 engagement especially during the MCOs. The research team contacted participants to  
489 schedule home visits for sample collection to ensure a higher retention rate. Overall, from  
490 our observation, keeping participants informed about the study's progress, providing timely  
491 updates, and addressing their concerns were the key factors in building strong trusting  
492 relationships and rapport with the participants. These strategies strengthened our  
493 enrollment and retention efforts, fostering an engaged cohort. This approach has not only  
494 empowered individuals from various backgrounds to participate actively but also  
495 contributed to the overall success and validity of our research study.

496

### 497 **5.2 Modification of the study designs**

498

499 A limitation within our study design arises from the early conceptualization during the initial  
500 stages of the pandemic, resulting in cross-exposure between the initial study groups.  
501 Consequently, the distinction between breakthrough infections in the vaccination-only  
502 (VAC) group and vaccinations in the infection-only (COV) group was not effectively  
503 addressed. This limitation hindered the ability to exclusively isolate the immune response



504 from either VAC or COV throughout the entire blood sampling period. Furthermore, there is  
505 a possibility of underreported data on previous infection history due to the potential  
506 occurrence of asymptomatic infections. Another important limitation is that we only  
507 gathered data on symptoms and adverse events following immunization (AEFI) during the  
508 enrollment phase, without continuing to track during the breakthrough infections and  
509 reinfections in the later stages of the study.

510

511 We acknowledged the potential confounding effects from individual behaviors and  
512 personality traits when assessing the influence of vaccination timing on SARS-CoV-2  
513 breakthrough infection risk. Variability in risk perception and health-related behaviors  
514 between early vaccine recipients and those delaying vaccination, alongside socioeconomic  
515 disparities and differential access to healthcare, present significant challenges in analysis.  
516 Therefore, we will perform multivariate analyses and sensitivity analyses to assess the  
517 robustness of findings across different exposures, while using regression modeling control  
518 for confounding variables.

519

### 520 **5.3 Cross-exposure of SARS-CoV-2 vaccination and infection from initial groupings**

521

522 With the emergence of breakthrough infections and reinfections, we reconsidered the initial  
523 stratification of VAC and COV groups due to cross-exposure between the 2 groups.  
524 Supplemental Figure 3 shows an overlapped graph depicting the sampling period of  
525 LONGTONG-SARS2 study and the national SARS-CoV-2 variants genomic surveillance data in  
526 Malaysia. We expected significant hybrid immunity, arising from both natural infections  
527 followed by vaccination and vice-versa, to be present in both participant groups amidst the  
528 emergence of highly transmissible variants and the high number of COVID-19 cases in  
529 combination with mass country-wide vaccination campaigns (Supplemental Figure 3,  
530 Supplemental Figure 4, Supplemental Table 2).

531

532 During the peak of Delta and Omicron circulation, when these variants constituted 75% of  
533 the COVID-19 cases in Malaysia (Hodcroft, 2023), we enrolled 67 and 222 participants,  
534 respectively. Those participants were likely infected with those variants based on the  
535 dominant circulating variants during the enrollment (Supplemental Figure 4). However, the

536 present study did not involve sequencing respiratory samples. Instead, we inferred the  
537 participants' likely infection variant based on contemporary national genomic sequencing  
538 data [25].

539

540 During the peak of the Delta variant, 16.8% of all participants (n=67/400) were identified,  
541 with 23.6% (n=51/216) from the COV group and 8.7% (n=16/184) from the VAC group.  
542 Similarly, during the peak of the Omicron variant, 55.5% of all participants (n=222/400), with  
543 67.1% (n=145/216) in the COV group and 41.8% (n=77/184) in the VAC group.

544

545 Cross-exposure of SARS-CoV-2 vaccination and infection was observed from the initially  
546 stratified participants throughout the 12-month sampling period in the present study  
547 (Supplemental Table 2). Almost two-thirds (68.8%) of all study participants (n=275/400)  
548 encountered breakthrough infections. Notably, within the VAC group, more than half  
549 (53.3%) experienced breakthrough infections (n=98/184). Furthermore, almost every  
550 participant (98.6%) in the COV group (n=213/216) had received a full primary series COVID-  
551 19 vaccination due to mass vaccination campaigns implemented by the Malaysian  
552 government to ensure a high vaccination rate. Detailed demographic data of this subset  
553 population can be referred to the Table 1 and Table 2.

554

#### 555 **5.4 Participants' demographics and study groups restructuring**

556

557 Of the 400 participants, 70% (n=280) completed the blood sampling follow-ups  
558 (Supplemental Table 1). Among them, 66.8% (n=267) completed the follow-ups and CRF.  
559 These 267 were included for the final analyses.

560

561 Of the 267 participants, 52.8% (n=141) are from the VAC group, while the remaining 47.2%  
562 (n=126) belonged to the COV group (Table 1). Demographic characteristics revealed a  
563 balanced sex distribution, with 39.3% male and 60.7% female participants (p=0.253). Age  
564 distribution showcased a significant difference across age groups (p=0.001), with 20.2%,  
565 47.6%, and 32.2% of participants falling into age groups  $\leq 30$  years old, 31-40 years old, and  
566  $> 41$  years old, respectively. Median age of participants was 36 years old ( $\bar{x}$ =37, IQR=9 to  
567 73).

568

569 **Table 1.** Distribution of sociodemographic factors by sex and age groups according to initial participant  
 570 grouping, VAC versus COV. Abbreviations: VAC = SARS-CoV-2-naïve prior to scheduled BNT162b2 primary  
 571 series vaccination; COV = recently recovered from COVID-19 infection.

	VAC	COV	Total	X <sup>2</sup> Stat (df)	p-value
<b>Sex</b>					
Male	60 (42.5%)	45 (35.7%)	105 (39.3%)		
Female	81 (57.4%)	81 (64.3%)	162 (60.7%)	1.30 (1)	0.253
<b>Age group</b>					
≤ 30	17 (12.1%)	37 (29.4%)	54 (20.2%)		
31-40	68 (48.2%)	59 (46.8%)	127 (47.6%)	15.11 (2)	0.001
> 40	56 (39.7%)	30 (23.8%)	86 (32.2%)		
<b>Total</b>	<b>141 (52.8%)</b>	<b>126 (47.2%)</b>	<b>267 (100%)</b>		

572

573 In the revised study participant grouping, our study delineated five distinct subgroups.  
 574 These included “E/VAC(BNT)/COVID-ve”, “E/VAC(BNT)/COVID+ve”,  
 575 “VAC(BNT)/COVID+ve/E”, “VAC(any)/COVID+ve/E”, and “COVID+ve/E/VAC(any)” (Figure 3).  
 576 Where the acronym is defined as “E” is Enrolled; “VAC(\*\*\*)” is vaccination type either  
 577 BNT162b2 or any vaccinations; “COVID\*ve” is the status of COVID-19 infection either  
 578 positive or negative. Forward slashes in the revised group names separate distinct events in  
 579 chronological order.

580

581 In detail, the five distinct groups within our study cohort, each defined by their vaccination  
 582 history and COVID-19 infection status:

- 583 1. "E/VAC(BNT)/COVID-ve" = this group comprised participants who were enrolled,  
 584 received the primary series of BNT162b2 vaccination, and did not experience a  
 585 breakthrough COVID-19 infection.
- 586 2. "E/VAC(BNT)/COVID+ve" = this group comprised participants who were enrolled,  
 587 received the primary series of BNT162b2 vaccination, and had documented  
 588 breakthrough COVID-19 infection.

- 589 3. "VAC(BNT)/COVID+ve/E" = this group comprised participants who received the  
 590 primary series of BNT162b2 vaccination prior to enrollment in the study, later had  
 591 documented breakthrough COVID-19 infection, and subsequently enrolled the study.  
 592 4. "VAC(any)/COVID+ve/E" = this group comprised participants vaccinated with a  
 593 primary series of COVID-19 vaccination (other than BNT162b2), later had  
 594 documented breakthrough COVID-19 infection, and subsequently enrolled in the  
 595 study.  
 596 5. "COVID+ve/E/VAC(any)" = this group comprised participants who had documented  
 597 COVID-19 infection prior to study enrollment and subsequently received a primary  
 598 series of COVID-19 vaccination (other than BNT162b2).  
 599

600 The results of the Chi-square test indicate that the distribution of sex and age groups in the  
 601 revised study group is comparable to that in the initial study groups (Table 2). Median age of  
 602 participants in the revised group was 36 years old ( $\bar{x}$ =38, IQR=15 to 65). There were no  
 603 significant differences in sex ( $p$ -value=0.324); however, a significant difference was  
 604 observed among age groups ( $p$ -value=0.001).  
 605

606 **Table 2.** Distribution of sociodemographic factors of sex and age group by revised study group.

	E/VAC(BNT)/COVID-ve	E/VAC(BNT)/COVID+ve	VAC(BNT)/COVID+ve/E	VAC(any)/COVID+ve/E	COVID+ve/E/VAC(any)	Total	X <sup>2</sup> Stat (df)	P-value
<b>Sex</b>								
Male	27 (26%)	30 (29%)	22 (21%)	14 (13%)	12 (11%)	105 (39%)	4.66 (4)	0.324
Female	25 (15%)	55 (34%)	40 (25%)	25 (15%)	17 (11%)	162 (61%)		
<b>Age group</b>								
<=30	7 (13%)	10 (19%)	12 (22%)	18 (33%)	7 (13%)	54 (20%)	26.94 (8)	0.001
31-40	24 (19%)	42 (33%)	35 (28%)	15 (12%)	11 (9%)	127 (48%)		
>40	21 (24%)	33 (38%)	15 (17%)	6 (7%)	11 (13%)	86 (32%)		
<b>Total</b>	<b>52 (20%)</b>	<b>85 (32%)</b>	<b>62 (23%)</b>	<b>39 (15%)</b>	<b>29 (11%)</b>	<b>267 (100%)</b>		

607  
 608 Additionally, for the cross-sectional PBMCs collection, we will conduct a comparative  
 609 analysis of cellular immune responses among individuals in both homologous and  
 610 heterologous vaccination groups, further stratifying them based on breakthrough infection  
 611 occurrence (Supplemental Table 3). The cellular responses within these vaccinated groups  
 612 will be compared with those of an unvaccinated control group.  
 613

## 614 6.0 FUTURE DIRECTIONS

615

616 Studies are planned or ongoing examining the kinetics of antibody production post-SARS-  
617 CoV-2 vaccination and infection. Preliminary analysis indicates variations among the revised  
618 study groups, prompting further investigation into the factors influencing the rate of decline  
619 and persistence of antibody levels over time in response to breakthrough infection and/or  
620 reinfection. Furthermore, in our exploration of the potential correlates of protection, we will  
621 examine the specific antibody, memory B cell, and proteomic response that best correlate  
622 with durability and functionality of immune response against SARS-CoV-2. These analyses  
623 are pending. Additional investigations will examine how previous exposures to SARS-CoV-2,  
624 whether through vaccination or infection, impact adaptive immune responses during  
625 subsequent exposures to elucidate the role of memory B cells and memory T cells in  
626 sustaining long-term immunity. Taken together, the cohort from the present study will be  
627 leveraged to delve into a focused exploration of the heterogeneity of SARS-CoV-2 immune  
628 responses arising from diverse vaccination regimens and antigen exposures.

629

## 630 **7.0 ETHICS, COLLABORATION, AND DISSEMINATION**

631

632 This study is strictly adhered to the ethical principles outlined in the World Medical  
633 Association Declaration of Helsinki. Ethical approval was obtained from Universiti Malaya  
634 Medical Centre (MREC-UMMC SID: 2021226-9886), ensuring compliance with all relevant  
635 federal regulations governing the protection of human subjects. All study personnel hold  
636 certifications in Good Clinical Practice. The study is supported by the U.S. Naval Medical  
637 Research Unit INDO PACIFIC (NAMRU INDO PACIFIC) and will adhere to the NAMRU INDO  
638 PACIFIC regulations for publication in open-access journals and data management.

639

640 The research team welcomes opportunities for potential research collaborations. The data  
641 from this study can be obtained upon request from the corresponding author, SAB.  
642 Permission to access data and analytical files is contingent upon approval from the relevant  
643 research ethics committees and the data custodian. The LONGTONG-SARS2 biological  
644 materials and data custodians are the Tropical Infectious Diseases Research and Education  
645 Centre (TIDREC) and the U.S. Naval Medical Research Unit INDO PACIFIC (NAMRU INDO  
646 PACIFIC).

647

648 Key study findings will be disseminated through various channels to ensure widespread and  
649 transparent communication. Results will be published in journals for the scientific  
650 community. Furthermore, the findings will be presented at national and international  
651 conferences, facilitating knowledge exchange and fostering collaboration with researchers  
652 worldwide. Notably, the study team is committed to communicating the results to the study  
653 population. By ensuring direct and accessible communication, participants will be informed  
654 on the progress and outcomes of the present study in which they have actively participated.  
655

### 656 **Funding**

657 The study was supported by funds from the U.S. Joint Program Executive Office under the  
658 2020-2021 Coronavirus Aid, Relief, and Economic Security (CARES) Act.

659

### 660 **Acknowledgment**

661 We acknowledge the dedication and commitment of the voluntary study participants  
662 without whom this project would not have been possible.

663

### 664 **Authors' contribution**

665 SAB had full access to all the study data and takes responsibility for the integrity of the data  
666 and the accuracy of the data analysis. Study concept and design: SAB, JJ, RH, NCK.  
667 Participant enrollment and data collection: NCK, JJ. Sample processing: NCK. Laboratory  
668 work: NCK. Data analysis and interpretation: NCK. Original drafting of manuscript: NCK.  
669 Editing and revision of manuscript: NCK, JJ, HY, HN, AL, RH, SAB. Study coordination: SAB, JJ,  
670 HY, HN, AL, RH, NCK. All authors read and approved the final article.

671

672 **Disclaimer.** The views expressed in this article are those of the authors and should not be  
673 construed to reflect the official policy or represent the positions of the U.S. Department of  
674 the Navy, Department of Defense, nor the United States Government. LT Huy C. Nguyen,  
675 MC, USN NAMRU-IP; LCDR Robert D. Hontz, MSC, NAMRU-IP; CAPT Andrew G. Letizia, MC,  
676 USN NAMRU-IP are military service members. This work was prepared as part of their  
677 official duties. Title 17 U.S.C. 105 provides that 'copyright protection under this title is not  
678 available for any work of the United States Government.' Title 17 U.S.C. 101 defines a U.S.

679 Government work as work prepared by a military service member or employee of the U.S.

680 Government as part of that person's official duties.

681

## 682 **Competing interests**

683 The authors declare that there is no competing interest.

684

## 685 **References**

686

- 687 1. Kaufer AM, Theis T, Lau KA, et al. Laboratory biosafety measures involving SARS-CoV-  
688 2 and the classification as a Risk Group 3 biological agent. *Pathology*. 2020;52: 790.  
689 doi:10.1016/J.PATHOL.2020.09.006
- 690 2. World Health Organization. WHO COVID-19 dashboard. 2024 [cited 20 Mar 2024].  
691 Available: <https://data.who.int/dashboards/covid19/cases?n=c>
- 692 3. Alimohamadi Y, Tola HH, Abbasi-Ghahramanloo A, et al. Case fatality rate of COVID-  
693 19: a systematic review and meta-analysis. *J Prev Med Hyg*. 2021;62: E311–E320.  
694 doi:10.15167/2421-4248/JPMH2021.62.2.1627
- 695 4. Dugan HL, Stamper CT, Li L, et al. Profiling B cell immunodominance after SARS-CoV-2  
696 infection reveals antibody evolution to non-neutralizing viral targets. *Immunity*.  
697 2021;54: 1290-1303.e7. doi:10.1016/J.IMMUNI.2021.05.001
- 698 5. Jeffery-Smith A, Burton AR, Lens S, et al. SARS-CoV-2-specific memory B cells can  
699 persist in the elderly who have lost detectable neutralizing antibodies. *J Clin Invest*.  
700 2022;132. doi:10.1172/JCI152042
- 701 6. Liu Y, Zeng Q, Deng C, et al. Robust induction of B cell and T cell responses by a third  
702 dose of inactivated SARS-CoV-2 vaccine. *Cell Discov*. 2022;8. doi:10.1038/S41421-  
703 022-00373-7
- 704 7. Ortega N, Ribes M, Vidal M, et al. Seven-month kinetics of SARS-CoV-2 antibodies and  
705 role of pre-existing antibodies to human coronaviruses. *Nat Commun*. 2021;12.  
706 doi:10.1038/S41467-021-24979-9
- 707 8. Seow J, Graham C, Merrick B, et al. Longitudinal observation and decline of  
708 neutralizing antibody responses in the three months following SARS-CoV-2 infection  
709 in humans. *Nat Microbiol*. 2020;5: 1598–1607. doi:10.1038/S41564-020-00813-8
- 710 9. Gerhards C, Thiaucourt M, Kittel M, et al. Longitudinal assessment of anti-SARS-CoV-2  
711 antibody dynamics and clinical features following convalescence from a COVID-19  
712 infection. *Int J Infect Dis*. 2021;107: 221–227. doi:10.1016/J.IJID.2021.04.080
- 713 10. Guiomar R, Santos AJ, Melo AM, et al. Monitoring of SARS-CoV-2 Specific Antibodies  
714 after Vaccination. *Vaccines*. 2022;10. doi:10.3390/VACCINES10020154
- 715 11. Seki Y, Yoshihara Y, Nojima K, et al. Safety and immunogenicity of the Pfizer/BioNTech  
716 SARS-CoV-2 mRNA third booster vaccine dose against the BA.1 and BA.2 Omicron  
717 variants. *Med*. 2022;3: 406-421.e4. doi:10.1016/J.MEDJ.2022.04.013
- 718 12. Stephens DS, McElrath MJ. COVID-19 and the Path to Immunity. *JAMA*. 2020;324:  
719 1279–1281. doi:10.1001/JAMA.2020.16656
- 720 13. Castro-Dopico X, Ols S, Loré K, et al. Immunity to SARS-CoV-2 induced by infection or  
721 vaccination. *J Intern Med*. 2022;291: 32–50. doi:10.1111/JOIM.13372

- 722 14. Tarke A, Coelho CH, Zhang Z, et al. SARS-CoV-2 vaccination induces immunological T  
723 cell memory able to cross-recognize variants from Alpha to Omicron. *Cell*. 2022;185:  
724 847-859.e11. doi:10.1016/J.CELL.2022.01.015
- 725 15. Evans JP, Liu S-L. Challenges and Prospects in Developing Future SARS-CoV-2  
726 Vaccines: Overcoming Original Antigenic Sin and Inducing Broadly Neutralizing  
727 Antibodies. *J Immunol*. 2023;211: 1459–1467. doi:10.4049/JIMMUNOL.2300315
- 728 16. Wang Q, Guo Y, Tam AR, et al. Deep immunological imprinting due to the ancestral  
729 spike in the current bivalent COVID-19 vaccine. *Cell Rep Med*. 2023;4:  
730 doi:10.1016/J.XCRM.2023.101258
- 731 17. Costela-Ruiz VJ, Illescas-Montes R, Puerta-Puerta JM, et al. SARS-CoV-2 infection: The  
732 role of cytokines in COVID-19 disease. *Cytokine Growth Factor Rev*. 2020;54: 62.  
733 doi:10.1016/J.CYTOGFR.2020.06.001
- 734 18. Virolainen SJ, VonHandorf A, Viel KCMF, et al. Gene-environment interactions and  
735 their impact on human health. *Genes Immun*. 2023;24: 1–11. doi:10.1038/S41435-  
736 022-00192-6
- 737 19. MacIntyre CR, Costantino V, Trent M. Modelling of COVID-19 vaccination strategies  
738 and herd immunity, in scenarios of limited and full vaccine supply in NSW, Australia.  
739 *Vaccine*. 2022;40: 2506–2513. doi:10.1016/J.VACCINE.2021.04.042
- 740 20. Guo Y, Logan HL, Glueck DH, et al. Selecting a sample size for studies with repeated  
741 measures. *BMC Med Res Methodol*. 2013;13: 1–8. doi:10.1186/1471-2288-13-  
742 100/FIGURES/3
- 743 21. Hartley GE, Edwards ESJ, Aui PM, et al. Rapid generation of durable B cell memory to  
744 SARS-CoV-2 spike and nucleocapsid proteins in COVID-19 and convalescence. *Sci*  
745 *Immunol*. 2020;5. doi:10.1126/SCIIMMUNOL.ABF8891
- 746 22. Peng P, Deng H, Li Z, et al. Distinct immune responses in the early phase to natural  
747 SARS-CoV-2 infection or vaccination. *J Med Virol*. 2022;94: 5691–5701.  
748 doi:10.1002/JMV.28034
- 749 23. Li C, Yue L, Ju Y, et al. Serum Proteomic Analysis for New Types of Long-Term  
750 Persistent COVID-19 Patients in Wuhan. *Microbiol Spectr*. 2022;10:  
751 doi:10.1128/SPECTRUM.01270-22
- 752 24. Pedersen AB, Mikkelsen EM, Cronin-Fenton D, et al. Missing data and multiple  
753 imputation in clinical epidemiological research. *Clin Epidemiol*. 2017;9: 157.  
754 doi:10.2147/CLEP.S129785
- 755 25. Hodcroft E. Overview of variants in countries: Malaysia. 2023 [cited 20 December  
756 2023]. Available: <https://covariants.org/per-country?country=Malaysia>
- 757

758

## 759 **List of figures**

760

761 **Figure 1.** Schedule of longitudinal blood sampling collection for up to 12 months and the  
762 cross-sectional PBMCs sampling period during the expected waning antibody phases.

763



764 **Figure 2.** Overall LONGTONG-SARS2 study design of laboratory analysis comprising of i)  
765 serology monitoring, ii) B- and T- memory cell functionality, iii) B lymphocyte population  
766 counting, iv) cytokine profiling, and v) proteomic profiling.

767

768 **Figure 3.** Flow chart of group stratification changes from initial study groupings to revised  
769 study groupings to accommodate the cross-exposures in the participants immune response.

# BNT162b2 vaccination group (VAC)

medRxiv preprint doi: <https://doi.org/10.1101/2024.03.26.24304850>; this version posted March 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).



Pre-vaccination



Day-7  
post first dose



Day-14  
post first dose



Day-7  
post second dose



Day-14  
post second dose



Day-90  
post first dose



Day-180  
post first dose



Day-270  
post first dose



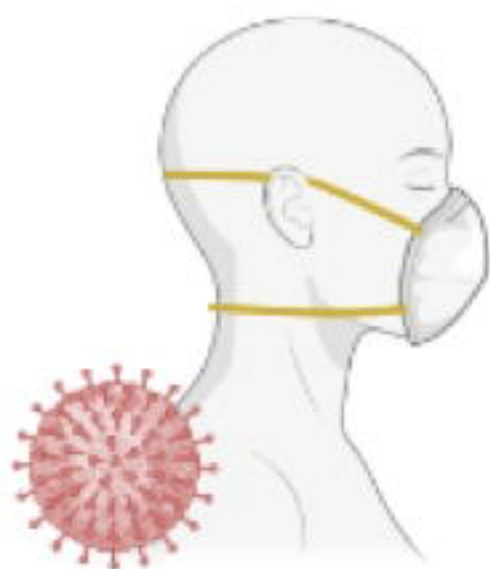
Day-360  
post first dose



Cross-sectional PBMCs sampling period



# COVID-19 infected group (COV)



1

Day-14  
post symptom  
onset



2

Day-30  
post symptom  
onset



3

Day-60  
post symptom  
onset



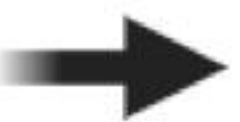
4

Day-90  
post symptom  
onset



5

Day-180  
post symptom  
onset



6

Day-270  
post symptom  
onset



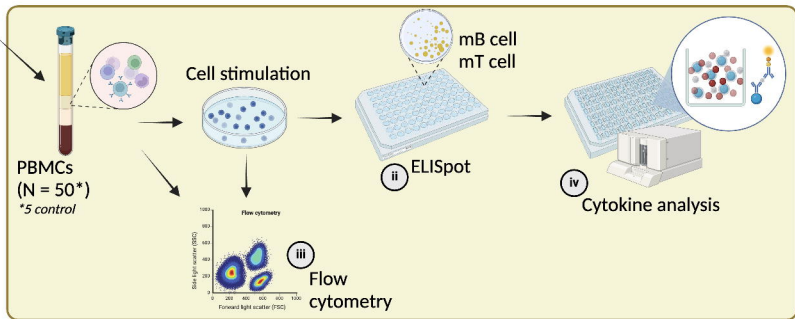
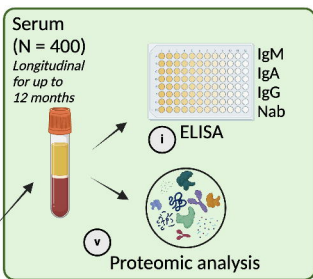
7

Day-360  
post symptom  
onset



Cross-sectional PBMCs sampling period





**Total LONGTONG participants during recruitment (N = 400)**

**Initial study groupings of LONGTONG in the early pandemic phase**  
with the objective to compare between those who were receiving BNT162b2 vaccination VS those who were infected with SARS-CoV-2

Screening inclusion and exclusion criteria

1. Aged 7 years or older
2. Healthy, no active SARS-CoV-2 during recruitment
3. Agreed to give informed consent
4. Not an immunosuppressive patient
5. Able to attend scheduled follow-ups

1. Receiving BNT162b2  
2. Negative for baseline SARS-CoV-2 anti-S IgG serology

**VAC group (n = 184)**

1. Infected with SARS-CoV-2 within 30 days  
2. Able to provide proof of infection either from rapid test kits or PCR

**COV group (n = 216)**

medRxiv preprint doi: <https://doi.org/10.1101/2024.03.26.24304850>; this version posted March 28, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Longitudinal blood sampling took place during the early phase of COVID-19 pandemic to the transition of endemic phases (March 2021 until May 2023)

Major limitations include:

1. Cross-exposure of vaccination and infection between initial groupings of LONGTONG.
2. Individuals receiving BNT162b2 were infected with SARS-CoV-2
3. Individuals infected with SARS-CoV-2 were mostly had received COVID-19 vaccinations.
4. Case report forms (CRFs) and questionnaires of COVID-19 severity, symptoms, and AEFI for post-vaccinations/boosters were only collected during the enrollment.
5. Movement control orders (MCOs) affecting the retention rate.

from  
*Total LONGTONG participants during enrollment*

Provide complete data on vaccinations/boosters/infection histories  
(n = 374)

Complete 12-month longitudinal blood sampling  
(n = 280)

Provide complete data on vaccinations/boosters/infection histories AND complete 12-month longitudinal blood sampling  
(n = 267)

**Revised study groupings of LONGTONG in the later pandemic phase**  
with the considerations of cross-exposures of SARS-CoV-2 vaccinations/boosters/infection

from  
*VAC group*

from  
*COV group*

No breakthrough infections (BTIs) throughout the 12-month follow-ups based on rapid test kits or PCR or 4-fold increase of SARS-CoV-2 IgG from previous blood sampling

**Group "E/VAC(BNT)/COVID-ve" (n = 52)**

Experienced breakthrough infections (BTIs) throughout the 12-month follow-ups based on rapid test kits or PCR or 4-fold increase of SARS-CoV-2 IgG from previous blood sampling

**Group "E/VAC(BNT)/COVID+ve" (n = 85)**

Received BNT162b2 vaccination before infected with SARS-CoV-2

**Group "VAC(BNT)/COVID+ve/E" (n = 62)**

Received vaccines other than BNT162b2 before infected with SARS-CoV-2

**Group "VAC(any)/COVID+ve/E" (n = 39)**

Had not received any COVID-19 vaccinations before infected with SARS-CoV-2

**Group "COVID+ve/E/VAC(any)" (n = 29)**

Cross-sectional PBMCs collection were drawn randomly from participants who agreed for cellular immunity study from Day-180 to Day-360 timepoints  
(n = 45)