

1 **TITLE**

2 Comparative analysis of symptom profile and risk of death associated with infection by SARS-CoV-2
3 and its variants in Hong Kong

4
5 **AUTHORS**

6
7 Kin On KWOK^{1,2,3,*,#}, Wan In WEI^{1,*}, Edward B. MCNEIL¹, Arthur TANG⁴, Julian Wei Tze TANG⁵,
8 Samuel Yeung Shan WONG^{1,@}, Eng Kiong YEOH^{6,@}

9
10 ¹ JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong
11 Special Administrative Region, China

12 ² Stanley Ho Centre for Emerging Infectious Diseases, The Chinese University of Hong Kong, Hong
13 Kong Special Administrative Region, China

14 ³ Hong Kong Institute of Asia-Pacific Studies, The Chinese University of Hong Kong, Hong Kong
15 Special Administrative Region, China

16 ⁴ School of Science, Engineering and Technology, RMIT University, Vietnam

17 ⁵ Respiratory Sciences, University of Leicester, Leicester, United Kingdom; Clinical Microbiology,
18 Leicester Royal Infirmary, Leicester, United Kingdom

19 ⁶ Centre for Health Systems and Policy Research, The Chinese University of Hong Kong, Hong Kong
20 Special Administrative Region, China

21 * Joint first author

22 @ Joint senior author

23 #Correspondence : kkokwok@cuhk.edu.hk

24
25 **KEYWORDS**

26 COVID-19; variants; symptom profile; symptomatology

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

27 ABSTRACT

28 **Introduction:** The recurrent multi-wave nature of COVID-19 necessitates updating its symptomatology.
29 Before the omicron era, Hong Kong was relatively unscathed and had a low vaccine uptake rate among
30 the old-old, giving us an opportunity to study the intrinsic severity of SARS-CoV-2 variants. A
31 comparison of symptom patterns across variants and vaccination status in Hong Kong has yet to be
32 undertaken. The intrinsic severity of variants and symptoms predictive of severe outcomes are also
33 understudied as COVID-19 evolves. We therefore aim to characterize the effect of variants on symptom
34 presentation, identify the symptoms predictive and protective of death, and quantify the effect of
35 vaccination on symptom development.

36 **Methods:** With the COVID-19 case series in Hong Kong from inception to 25 August 2022, an iterative
37 multi-tier text-matching algorithm was developed to identify symptoms from free text. Cases were fully
38 vaccinated if they completed two doses. Multivariate regression was used to measure associations
39 between variants, symptom development, death and vaccination status. A least absolute shrinkage and
40 selection operator technique was used to identify a parsimonious set of symptoms jointly associated with
41 death.

42 **Results:** Overall, 70.9% (54450/76762) of cases were symptomatic. We identified a wide spectrum of
43 symptoms (n=102), with cough, fever, runny nose and sore throat being the most common (8.16-47.0%).
44 Intrinsically, the wild-type and delta variant caused similar symptoms, with runny nose, sore throat, itchy
45 throat and headache more frequent in the delta cohort; whereas symptoms were heterogeneous between
46 the wild-type and omicron variant, with seven symptoms (fatigue, fever, chest pain, runny nose, sputum
47 production, nausea/vomiting and sore throat) more frequent in the omicron cohort. With full vaccination,
48 omicron was still more likely than delta to cause fever. Fever, blocked nose and shortness of breath were
49 robustly jointly predictive of death as the virus evolved. Number of vaccine doses required for reduction
50 in occurrence varied by symptoms.

51 **Discussion:** This is the first large-scale study to evaluate the changing symptomatology by COVID-19
52 variants and vaccination status using free-text reporting by patients. We substantiate existing findings that

53 omicron has a different clinical presentation compared to previous variants. Syndromic surveillance can
54 be bettered with reduced reliance on symptom-based case identification, increased weighing on
55 symptoms robustly predictive of mortality in outcome prediction, strengthened infection control in care
56 homes through universal individual-based risk assessment to enable early risk stratification, adjusting the
57 stockpile of medicine to tally with the changing symptom profiles across vaccine doses, and
58 incorporating free-text symptom reporting by patients.

59

60 Word count = 404

61 INTRODUCTION

62

63 *Evolving SARS-CoV-2 epidemiology in Hong Kong*

64

65 Hong Kong has experienced five waves of Coronavirus Disease 2019 (COVID-19), caused by severe
66 acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1]. The first two waves were driven by the
67 ancestral wild-type virus, the third and the fourth were dominated by lineages B.1.1.63 and B.1.36.27,
68 respectively, and the early fifth wave was mainly caused by the omicron subvariant BA.2 [1]. Hong Kong
69 was relatively unscathed in the first four waves, which can be attributed to a hospital-based approach
70 under which all confirmed or suspected cases were isolated. Before the fifth wave, Hong Kong was
71 characterized by a low proportion (18%) of the old-old completing a two-dose primary vaccination series
72 and predominantly vaccine-induced immunity at the population level [2]. To cope with the surge in cases
73 during the fifth wave, an online declaration system was established for the community to self-report their
74 infection and symptoms [3]. Against this background, the disease progression in Hong Kong enables
75 revisiting the clinical spectrum of SARS-CoV-2 of primary infection episodes across vaccination groups
76 amid a changing landscape of viral pathogenicity [4].

77

78 *Importance of identifying the symptom profile of SARS-CoV-2*

79

80 The recurrent multi-wave nature of COVID-19 [5] makes it worth updating the symptomatology to
81 improve disease recognition, particularly when it involves the telltale signs of severe health outcomes. As
82 of 26 March 2023, weekly hospitalizations by COVID-19 exceeded that of influenza by 13-fold in the
83 United Kingdom [6]. Though symptoms triggered by respiratory illnesses are generally alike [7], the
84 persistent clinical sequelae after the acute phase of infection [8] and the complexity of symptoms added
85 by the omicron variant [9] highlight the need to differentiate COVID-19 from other respiratory diseases.
86 While the circulating omicron variant may induce less severe symptoms than its preceding counterparts
87 [10], this benignity can be counterbalanced by an increased incidence of infection disrupting society via

88 sickness. In fact, 46% of the patients infected with omicron regarded their illnesses as moderate or severe
89 [11].

90

91 *Research gaps*

92

93 Despite the importance of updating symptom profiles and disease burden as SARS-CoV-2 evolves, there
94 are **four** research gaps that need addressing.

95

96 **First**, symptom profiles of COVID-19 vary across regions [12], but to our knowledge, we are not aware
97 of any study comparing symptom patterns of infection across variants and vaccination status in Hong
98 Kong. Such a clinical spectrum was studied initially [13,14], but the varying pathogenicity of SARS-
99 CoV-2 [4] highlights the need to revisit the clinical presentation. Even under the same viral condition,
100 symptom profiles vary with the host properties and the immunity level of the affected population, which
101 is an interplay between natural infection and vaccine regimes (for example, the type of vaccine
102 administered and timing of vaccination).

103

104 **Second**, the intrinsic severity of SARS-CoV-2 variants are understudied and most published studies have
105 returned mixed findings. Intrinsic-severity estimates are indispensable information to plan upcoming
106 infection control strategies against the public burden caused by immune-evasive pathogens,
107 compromising the effectiveness of vaccination. On one hand, the intrinsic severity is less for omicron
108 according to Nyberg and colleagues who reported a larger reduction (comparing omicron with delta) in
109 the risk of hospitalization and death in unvaccinated cases than all cases [15], and by Zhang and
110 colleagues who reported no significant difference in the blood profiles of the fully vaccinated and the
111 non/partially vaccinated [16]. On the other hand, Whitaker and colleagues found that omicron cases were
112 more likely to report cold-like symptoms [17]. Therefore, it is important to unravel the intrinsic severity
113 of the concurrent and recent predominant SARS-CoV-2 variant, especially when the mainstream vaccine

114 administered in Hong Kong has low efficacy against symptomatic infection of past strains [18], and when
115 it takes time for the bivalent vaccine to be taken by the majority of the community.

116
117 **Third**, symptoms predictive of severe health outcomes are understudied as COVID-19 evolves. To our
118 knowledge, such information was studied in the wild-type era by Vahey and colleagues who reported that
119 vomiting, dyspnea, altered mental status, dehydration, and wheezing were significantly associated with
120 hospitalization [19]. However, differences in replication and pathogenicity among SARS-CoV-2 variants
121 [4] may present varied symptom-based predictors of mortality. By identifying which symptoms are
122 predictive of severe outcomes, patients can be better triaged for medical procedures and prioritization of
123 testing.

124
125 **Fourth**, established work based on structured responses of a predefined scope of symptoms might
126 underestimate the prevalence and diversity of symptoms. For example, Vihta and colleagues recorded 12-
127 16 symptoms [20] versus 32 preset symptoms in the ZOE app [21]). In addition, Malden and colleagues
128 showed that free-text data of electronic medical records increased the identification of more cases for
129 each symptom by 29-64% [22].

130 131 *Study aims and significance*

132
133 In light of the above research gaps, this study makes use of free-text symptom data self-reported by cases
134 to (i) characterize the (intrinsic and post-vaccination, wherever applicable) effect of variants on symptom
135 presentation, (ii) identify the symptoms which are predictive and protective of death across variants, and
136 (iii) quantify the effect of vaccination on symptom development. Following the World Health
137 Organization's end of the "public health emergency of international concern" declaration for COVID-19
138 [23], future infection control will focus on elderly and high-risk groups such that further investigations
139 into the telltale signs of COVID-19 (as it evolves) will guide public health policies during patient triage.

METHODS

Data curation

We analyzed the case series of COVID-19 infections provided by the Centre for Health Protection (CHP) between 23 January 2020 and 25 August 2022 from which the following information was available: age, sex, chronic diseases, case classification (local or imported), date of case reporting to the CHP, symptoms experienced at the date of reporting, date of symptom onset, and vaccination date and type. The symptom data was initially solicited by nurse interviews, which was replaced with self-reporting on 26 February 2022.

Iterative process of symptom extraction from free-text data

Symptoms were recorded in the form of free text in addition to four hard-coded symptoms (fever, chest pain, shortness of breath and breathing difficulty). Through compiling frequent lay-person descriptions by cases, a list was created to map their responses to standardized variables created based on the national ambulatory medical care survey (NAMCS) [24]. Classification by NAMCS was based on patients' expression of their symptoms serving as the input to health services, which fits the current context of COVID-19 cases reporting their symptoms. To increase the accuracy of symptom identification, an iterative multi-tier text-matching algorithm was developed (**Table S1**). A sample of 200 randomly selected cases was manually reviewed for symptom extraction to assess the accuracy of this algorithm.

Classification of variants

Cases reported on or before 30 June 2020 were regarded as wild-type cases [1]. Starting from the third wave (1 July 2020), cases not sequenced for the determination of the variant or failed/had insufficient

167 viral load for sequencing were unclassified. For the remaining cases, the variant was determined based on
168 the results of whole genome sequencing by CHP upon positive results of polymerase chain reaction or
169 rapid antigen testing.

171 *Definition of vaccination status*

172
173 Cases were defined as unvaccinated if they did not receive any dose of vaccine on or before the symptom
174 onset date (for symptomatic individuals) or the report date (for asymptomatic individuals). Cases were
175 considered to complete a vaccine dose if that dose was received at least 14 days before the symptom
176 onset date (for symptomatic individuals) or the report date (for asymptomatic individuals). Cases were
177 defined as fully vaccinated if they completed two vaccine doses. A 14-day window period was assumed
178 by convention as the human immune system takes time to build up protection [25].

180 *Statistical analysis*

181
182 Cases with incomplete data were excluded from the analysis (**Figure 1**). Descriptive statistics, including
183 proportions with binomial confidence intervals, were used to summarize the data. Multivariate regression
184 models were fit to examine **(i)** the effect of variants on symptom development, **(ii)** the association of
185 symptoms on elderly deaths, and **(iii)** the effect of vaccination on symptom development. The intrinsic
186 severity of COVID-19 variants among elderly cases, measured as the occurrence of death within 28 days
187 of symptom onset, was compared with Cox's proportional hazards regression. All models were adjusted
188 for age, sex, and reporting delay, defined as the time elapsed between symptom onset and the report date
189 for symptomatic cases. When considering the overall effect regardless of variants, the analyses were also
190 adjusted for variants. Adjusted odds ratios (aOR) and hazard ratios with 95% confidence intervals (CI)
191 were presented visually.

193 We used a least absolute shrinkage and selection operator (LASSO) technique under the generalized
194 linear regression framework to identify parsimonious symptoms jointly associated with death. The *glmnet*
195 package in R was used to fit these models [26]. The value of the regularization parameter was determined
196 using *k*-fold cross validation. The performance of these regression models was evaluated using the area
197 under the receiver operating characteristic curve (AUC). All analyses were performed in R (version 4.2).

199 RESULTS

201 *Subject characteristics*

203 As of 25 August 2022, there were 1,479,393 COVID-19 cases reported in Hong Kong, among which
204 76,762 were included in the analysis with the majority being BA.2 cases (**Figure 1**). The reporting period
205 of these cases ranged from 23 January 2020 to 19 July 2022. The study population consisted largely of
206 adults aged 20-59 years (74.2%), followed by the elderly aged ≥ 60 years (15.3%) and young adults or
207 minors (10.5%) (**Table 1**). The variant-specific proportion of local cases differed, ranging from over 95%
208 for omicron (BA.2) and delta to 34.5% for wild type. All 1,175 wild-type cases were recorded on or
209 before 30 June 2020, while the variants were mostly recorded in 2022. All wild-type cases were
210 unvaccinated, but the circulation of its variants or sub-variants was amid an increasing level of vaccine-
211 induced population immunity.

213 *Prevalence of symptoms*

215 The majority (70.9%) of the cases were symptomatic (**Table 1**). Compared with the hard-coded
216 symptoms, extraction from free-text data identified an additional 5336/11343 (47.0%) cases with fever,
217 714/2173 (32.9%) cases with shortness of breath or breathing difficulty, and 68/1770 (3.8%) with chest
218 pain. Overall, a wide spectrum of 102 symptoms was identified, among which 30 symptoms had an
219 overall prevalence of at least 0.5% (**Table S3**). The free-text extraction showed that cases used varied

expressions to describe their symptoms in the throat (for example, dry throat, itchy throat, sore throat, throat discomfort, throat irritation, tonsillitis) and in the muscle pain category (for example, bodyache, back pain, low extremity pain, unspecified muscle pain) (**Table S2**). Regardless of variants and vaccination status, the most common symptoms were cough (range: 22.5-47.0%), fever (17.9-37.2%), runny nose (8.16-33.1%), sore throat (15.0-29.1%) and headache (4.96-14.6%) (**Figure 2**). Among the wild-type cases, three more symptoms stood out: disturbance of taste and/or smell (6.38%), shortness of breath (5.19%) and diarrhea (4.94%).

Effect of variant on symptom presentation

Effects of variant on symptom presentation were examined among the unvaccinated symptomatic cases (**Figure 3**), implicating the intrinsic severity of the variants in context. Omicron was more likely than wild-type to be associated with fatigue (aOR: 2.83 [95% CI: 1.53-5.21]), fever (aOR: 1.96 [1.69-2.27]), chest pain (aOR:7.21 [2.96, 17.6]), runny nose (aOR: 1.36 [1.09-1.69]), expectoration (aOR: 2.54 [1.62-3.99]), nausea/vomiting (aOR:5.13 [2.53-10.39]) and sore throat (aOR: 1.68 [1.41-2.00]), but less likely to be associated with malaise, blocked nose, cough, diarrhea, throat discomfort, taste/smell disturbances, and headache. On the other hand, the symptoms associated with delta and the wild-type were more alike in that delta was more likely to be associated with only four symptoms: runny nose (aOR: 4.15 [2.30-7.49]), sore throat (aOR: 1.94 [1.08-3.48]), itchy throat (aOR: 36.58 [3.20-418.78]) and headache (aOR: 2.41 [1.24-4.68]). Differences in symptom presentation were also noted between omicron and delta for fever, runny nose, and headache.

Effects of variant on symptom presentation were examined among the fully vaccinated symptomatic cases (**Figure 4**). Omicron was more likely than delta to be associated with fever (aOR:1.66 [1.08-2.57]), but less likely to be associated with malaise, cough, runny nose, dry throat, taste/smell disturbances, dizziness, chest discomfort, palpitations (not shown in Figure 4) and headache.

247 *Association between symptom presentation and death*

248

249 The overall case fatality rate (CFR) was 0.25% (195/76,762), and that among the unvaccinated and the
250 fully vaccinated were 1.11% and 0.05% respectively. Intrinsically, the hazard of death (within 28 days of
251 symptoms onset) due to delta and omicron infections among elderly were similar, and were both higher
252 than that attributable to the ancestral wild-type (**Figure S1**). Among 75 decedents who were unvaccinated
253 and symptomatic, 18 symptoms were ever reported: fever (64.0%) being the most common, followed by
254 cough (37.3%), shortness of breath (25.3%) and sore throat (10.7%) (**Table 2**). Fever (aOR: 2.04 [1.22,
255 3.47]) and shortness of breath (aOR:2.25 [1.21, 4.03]) were predictive of death; whereas cough
256 (aOR:0.51 [0.30, 0.84]) and sore throat (aOR:0.43 [0.19, 0.87]) were protective.

257

258 Older age (≥ 60 years) significantly increased the odds of death for all symptoms (not shown in Table 2).
259 With most decedents being elderly, we searched for a robust set of symptoms which were jointly
260 predictive of death (**Table 3**). Overall, abdominal pain, blocked nose, malaise, shortness of breath, and
261 fever were predictive of death while protective symptoms included cough, diarrhea, chest pain, voice
262 disorders, sore throat, and loss of appetite. Among omicron cases, the predictive set included only
263 blocked nose, shortness of breath, and fever. The predictive power of these sets of symptoms was good
264 (AUC: 0.926).

265

266 *Effect of vaccination on symptom presentation*

267

268 Considering the difference in symptom presentation between the young (<60 years) and the elderly (≥ 60
269 years) (**Figure S2**), and older age is a significant risk factor for death, the extent of vaccination
270 alleviating symptom presentation was examined among symptomatic elderly cases (**Figure 5**). Receipt of
271 vaccination did not necessarily reduce the occurrence of symptoms, and in fact was associated with an
272 increased chance of developing sore throat and some respiratory symptoms such as blocked nose, cough,
273 and runny nose. The number of doses required to reduce individual symptoms also varied, from one dose

274 for reducing fever, to two doses for reducing shortness of breath, to three doses for reducing chills. We
275 also explored the effect of time elapsed since last vaccination on the occurrence of symptoms predictive
276 of death (**Figure 6**). Compared with the number of doses, time elapsed since vaccination given the same
277 number of doses did not alter the chance of occurrence of the predictive symptoms.

279 **DISCUSSION**

281 *Principal findings*

283 This is the first large-scale study to evaluate variations in symptom presentation by COVID-19 variants
284 and vaccination status using free-text reporting by patients. We substantiate existing findings that
285 omicron has a different clinical presentation compared to previous variants [27]. We identified a wide
286 spectrum of symptoms (n=102) associated with SARS-CoV-2 infection, with the most common ones
287 being cough, fever, runny nose and sore throat (8.16-47.0%). Intrinsically, symptoms caused by the wild-
288 type and delta were generally similar, except that runny nose, sore throat, itchy throat and headache were
289 more frequent in the delta cohort; whereas symptoms were heterogeneous between the wild-type and
290 omicron, with seven symptoms (fatigue, fever, chest pain, runny nose, sputum production,
291 nausea/vomiting and sore throat) more likely to be experienced by the omicron cohort, and seven
292 symptoms (malaise, blocked nose, cough, diarrhea, throat discomfort, taste/smell disturbances, and
293 headache) less likely to occur. Under full vaccination coverage, omicron was still more likely than delta
294 to cause fever. Furthermore, fever, blocked nose and shortness of breath were robustly predictive of death
295 even when the wild-type virus evolved to the omicron variant. Though the death rate was much lower in
296 the fully vaccinated cohort, vaccination does not necessarily reduce the occurrence of symptoms, and the
297 number of doses required to reduce symptoms varied by individual symptoms.

299 *Result implications*

301 Our findings have five public health implications.

302

303 **First**, the weighing of symptom-based predictions of COVID-19 should be reduced. We found a high

304 prevalence of upper respiratory tract infections among the predominant omicron cases in wave 5 with

305 differences by vaccination status. This result is consistent with five other studies [9,17,20,21,28].

306 Furthermore, we found a much lower incidence of “loss of taste and smell”, which, in the past, was used

307 as a test-triggering symptom [29,30], in the omicron cohort. These altogether increase the difficulty to

308 distinguish COVID-19 with symptom-based testing algorithms. In fact, Fragaszy and colleagues

309 calculated the specificity of such algorithms was as low as 47%-27% [31]. Although symptom-based

310 information is of reduced value to diagnosis, we urge updating the symptom variation in public health

311 messaging as failure to do so will hinder public awareness.

312

313 **Second**, the omicron variant may not be intrinsically less severe than preceding variants such that

314 continued surveillance among vulnerable populations should remain in force. Unlike regions with a high

315 incidence of natural infection prior to the omicron era, such as the United States [32], Hong Kong was

316 almost free of infection in the community [13,14] coupled with low vaccine uptake rate among the old-

317 old [2]. This allowed us to study the intrinsic severity of mortality among this disproportionately affected

318 group. In contrast to Nyberg and colleagues who reported a larger reduction in the risk of death (from

319 omicron to delta) among unvaccinated cases [15], we identified that the risk of death following SARS-

320 CoV-2 infection among the unvaccinated elderly was similar between omicron and delta (Table S1). We

321 also found that omicron was more likely to cause fever than delta among unvaccinated symptomatic cases

322 (Table S3, Figure 2). The reasons for the aforementioned contrasting findings need further investigation.

323 In addition, when compared with other co-circulating respiratory illnesses, omicron, especially the BA.2

324 sublineage, is not intrinsically milder and can be more neuropathogenic [33].

325

326 **Third**, symptom-based predictions for death are possible and can help refine infection control policies to

327 tackle COVID-19 in residential care homes for the elderly (RCHes). Though a major surge in severe

328 illnesses in RCHEs may be unlikely under a high vaccine or infection coverage [34], the delay in seeking
329 medical care among the elderly [35], the multi-fold increase in mortality among institutionalized elderly
330 [36], and the highest number of incidents attributable to RCHEs in the continual recurrent epidemics [6]
331 altogether suggest active syndromic surveillance is needed in RCHEs. Further to the proposed frailty
332 screening to be performed on every aged COVID-19 case to enable early risk stratification [37] and the
333 inclusion of certain underlying health conditions as risk characteristics (such as hypertension [38],
334 dyslipidemia and being bed-bound [39]), symptoms consistently predictive of severe clinical outcomes
335 can be added to this individual-based risk assessment. Further to Ryan and colleagues who showed that
336 severe COVID-19 could be predicted by symptoms in the wild-type era [40], we found certain symptoms
337 (fever and shortness of breath) were robustly predictive of death during the wild-type-to-omicron
338 transition. This finding is consistent with earlier reports in the wild-type era that shortness of breath was
339 predictive of death [19,39].

340

341 **Fourth**, the stockpile of genus epidemicus should be optimized with the changing symptom profiles
342 across vaccine doses given the long-term battle against COVID-19. In line with existing findings which
343 reported changing symptom profiles across vaccine doses [9,28], we found that fever was less common as
344 the number of vaccine doses increased, but some respiratory symptoms were still very common,
345 including cough, runny nose, blocked nose and sore throat. Though the effect of vaccines on reducing
346 severe outcomes following infection of COVID-19 is present across variants [41], it does not necessarily
347 lower the occurrence of symptoms with high transmissibility. In line with an early report that virus
348 shedding was initially highest in the upper respiratory tract [13], sore throat and runny nose were shown
349 to be associated with higher viral load following infection with SARS-CoV-2 [17]. This suggests that the
350 provision of medicine should be adjusted to suppress viral replication in human bodies with respect to the
351 changing symptom profiles, particularly in regions with intense social-mixing like Hong Kong where the
352 number of contacts reached 12.5 per day [42]. In Hong Kong, two antiviral agents (Molnupiravir and
353 Paxlovid) were available to improve clinical outcomes, but their effect on symptom alleviation needs
354 further research.

355

356 **Fifth**, future syndromic surveillance should be enriched by free-text reporting of symptoms by patients.

357 By applying a multi-tier text-matching algorithm on free-text symptom data, we identified a wider

358 disease spectrum, which exceeded the number of symptoms predefined by other studies [17,20,21].

359 Considering the four hard-coded symptoms in our dataset, we showed that a large proportion of

360 identification was contributed by additional free-text data (3.8-47%), which is in line with Malden and

361 colleagues who supplemented structured data of electronic medical records with natural language

362 processing of unstructured data and found 29-64% additional cases for each selected symptoms [22].

363

364 *Study strengths and limitations*

365

366 Clinical characteristics of SARS-CoV-2 were widely studied in the initial phase of the pandemic [43], but

367 comparisons between recent subtypes were sparse and the intrinsic severity of subsequent variants were

368 understudied. This is the first large-scale study which adds to the literature in this understudied area with

369 free-text symptom reporting data from the patient perspective. In addition, unlike existing studies, which

370 inferred the variant type to which a case belonged from the predominating strain circulating in the

371 community at that time [17], we defined our cohort based on confirmed laboratory results, minimizing

372 misclassification bias. However, there are **two** study limitations which bear mentioning. **First**, the

373 duration and level of symptoms were not recorded such that we were unable to provide a better

374 understanding of the difference between the same symptoms. **Second**, caution should be exercised when

375 interpreting differences in symptom across variants, which may be biased by the potential underreporting

376 of patient characteristics in the massive fifth omicron wave during which healthcare resources were

377 stringent and by the changing method of data collection over time.

378

379 *Conclusion*

380

381 Upon achieving herd immunity in the population to halt the exponential spread of SARS-CoV-2 [44], the
382 next item on the agenda is to handle the disease burden caused by the changing symptom profile. Existing
383 surveillance infrastructure should be updated with respect to changing symptomatology and mortality risk
384 as SARS-CoV-2 and the population immunity level evolve. Such infrastructure can be bettered with
385 reduced reliance on symptom-based case identification, increased weighing on symptoms robustly
386 predictive of mortality in outcome prediction, strengthened infection control in RCHes through universal
387 individual-based risk assessment to enable early risk stratification, and adjusting the stockpile of
388 medicine to target treatment arising from the changing clinical presentation against the background of
389 changing immunity in the population. In addition, real-time algorithms to incorporate patients' free-text
390 reporting of symptoms should be developed to enrich syndromic surveillance.

391

392 Word count = 3728

393

394 **ACKNOWLEDGEMENT**

395

396 We would like to acknowledge support from Health and Medical Research Fund (reference numbers:
397 INF-CUHK-1, 17160302, 18170312, CID-CUHK-A, COVID1903008), General Research Fund
398 (reference numbers: 14112818, 24104920), Wellcome Trust Fund (reference number: 200861/Z/16/Z)
399 and Group Research Scheme of The Chinese University of Hong Kong. This work formed part of the
400 doctoral thesis requirements for Wan In WEI.

401

402 **ETHICAL STATEMENT**

403 This study was approved by the Survey and Behavioral Research Ethics Committee of The Chinese
404 University of Hong Kong (reference: SBRE-19-595).

405

406 **DATA STATEMENT**

407 The availability of the data set is subject to approval from Hong Kong Centre for Health and Protection,
408 Department of Health and relevant government departments.

409

410 **COMPETING INTEREST**

411 None

412

413 **AUTHORS' CONTRIBUTIONS**

414 Conceptualized: Kin On KWOK (KOK), Wan In WEI (WIW)

415 Data curation: Edward B McNeil (EBM), WIW

416 Data analysis: EBM, WIW, KOK

417 First draft of the manuscript: WWI, KOK

418 Data interpretation: KOK, EBM, WIW

419 Manuscript editing: EK YEOH (EKY), Samuel YS WONG (SYSW), EBM, WWI, KOK, Arthur TANG

420 (AT), Julian TANG (JT)

421 Supervision: EKY, SYSW, KOK

422 Provision of critical comments: AT, SYSW, JT, EKY

423 REFERENCES

- 424 1. Wong S-C, Au AK-W, Lo JY-C, Ho P-L, Hung IF-N, To KK-W, Yuen K-Y, Cheng VC-C.
425 Evolution and Control of COVID-19 Epidemic in Hong Kong. *Viruses* 2022 Nov 14;14(11).
426 PMID:36423128
- 427 2. Smith DJ, Hakim AJ, Leung GM, Xu W, Schluter WW, Novak RT, Marston B, Hersh BS. COVID-
428 19 Mortality and Vaccine Coverage - Hong Kong Special Administrative Region, China, January 6,
429 2022-March 21, 2022. *MMWR Morb Mortal Wkly Rep* 2022 Apr 15;71(15):545–548.
430 PMID:35421076
- 431 3. HKSAR Government. “Declaration System for Individuals Tested Positive for COVID-19 Using
432 Rapid Antigen Test” launched. 2023. Available from:
433 <https://www.info.gov.hk/gia/general/202203/07/P2022030700768p.htm> [accessed Apr 7, 2023]
- 434 4. Shuai H, Chan JF-W, Hu B, Chai Y, Yuen TT-T, Yin F, Huang X, Yoon C, Hu J-C, Liu H, Shi J,
435 Liu Y, Zhu T, Zhang J, Hou Y, Wang Y, Lu L, Cai J-P, Zhang AJ, Zhou J, Yuan S, Brindley MA,
436 Zhang B-Z, Huang J-D, To KK-W, Yuen K-Y, Chu H. Attenuated replication and pathogenicity of
437 SARS-CoV-2 B.1.1.529 Omicron. *Nature* 2022 Mar;603(7902):693–699. PMID:35062016
- 438 5. Dhaliwal M, Small R, Webb D, Cluver L, Ibrahim M, Bok L, Nascimento C, Wang C, Garagic A,
439 Jensen L. Covid-19 as a long multiwave event: implications for responses to safeguard younger
440 generations. *BMJ* 2022 Jan 27;376:e068123. PMID:35086910
- 441 6. UK Health Security Agency. Weekly national Influenza and COVID-19 surveillance report Week 13
442 report (up to week 12 data). 2023 Mar. Available from:
443 [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/11](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1147550/Weekly_Flu_and_COVID-19_report_w13_v2.pdf)
444 [47550/Weekly_Flu_and_COVID-19_report_w13_v2.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1147550/Weekly_Flu_and_COVID-19_report_w13_v2.pdf)
- 445 7. Nypaver C, Dehlinger C, Carter C. Influenza and Influenza Vaccine: A Review. *J Midwifery*
446 *Womens Health* 2021 Jan;66(1):45–53. PMID:33522695
- 447 8. Cohen K, Ren S, Heath K, Dasmariñas MC, Jubilo KG, Guo Y, Lipsitch M, Daugherty SE. Risk of
448 persistent and new clinical sequelae among adults aged 65 years and older during the post-acute
449 phase of SARS-CoV-2 infection: retrospective cohort study. *BMJ* 2022 Feb 9;376:e068414.
450 PMID:35140117
- 451 9. Wang M, Liu Z, Wang Z, Li K, Tian Y, Lu W, Hong J, Peng X, Shi J, Zhang Z, Mei G. Clinical
452 characteristics of 1139 mild cases of the SARS-CoV-2 Omicron variant infected patients in
453 Shanghai. *J Med Virol* 2023 Jan;95(1):e28224. PMID:36238984
- 454 10. Ren S-Y, Wang W-B, Gao R-D, Zhou A-M. Omicron variant (B.1.1.529) of SARS-CoV-2:
455 Mutation, infectivity, transmission, and vaccine resistance. *World J Clin Cases* 2022 Jan 7;10(1):1–
456 11. PMID:35071500
- 457 11. Petersen MS, Í Kongsstovu S, Eliassen EH, Larsen S, Hansen JL, Vest N, Dahl MM, Christiansen
458 DH, Møller LF, Kristiansen MF. Clinical characteristics of the Omicron variant - results from a
459 Nationwide Symptoms Survey in the Faroe Islands. *Int J Infect Dis* 2022 Sep;122:636–643.
460 PMID:35811082
- 461 12. Kadirvelu B, Burcea G, Quint JK, Costelloe CE, Faisal AA. Variation in global COVID-19
462 symptoms by geography and by chronic disease: A global survey using the COVID-19 Symptom
463 Mapper. *EClinicalMedicine* 2022 Mar;45:101317. PMID:35265823
- 464 13. Kwok KO, Huang Y, Tsoi MTF, Tang A, Wong SYS, Wei WI, Hui DSC. Epidemiology, clinical

- 465 spectrum, viral kinetics and impact of COVID-19 in the Asia-Pacific region. *Respirology* 2021
466 Apr;26(4):322–333. PMID:33690946
- 467 14. Kwok KO, Wei WI, Huang Y, Kam KM, Chan EYY, Riley S, Chan HHH, Hui DSC, Wong SYS,
468 Yeoh EK. Evolving Epidemiological Characteristics of COVID-19 in Hong Kong From January to
469 August 2020: Retrospective Study. *J Med Internet Res* 2021 Apr 16;23(4):e26645. PMID:33750740
- 470 15. Nyberg T, Ferguson NM, Nash SG, Webster HH, Flaxman S, Andrews N, Hinsley W, Bernal JL,
471 Kall M, Bhatt S, Blomquist P, Zaidi A, Volz E, Aziz NA, Harman K, Funk S, Abbott S, COVID-19
472 Genomics UK (COG-UK) consortium, Hope R, Charlett A, Chand M, Ghani AC, Seaman SR,
473 Dabrera G, De Angelis D, Presanis AM, Thelwall S. Comparative analysis of the risks of
474 hospitalisation and death associated with SARS-CoV-2 omicron (B.1.1.529) and delta (B.1.617.2)
475 variants in England: a cohort study. *Lancet* 2022 Apr 2;399(10332):1303–1312. PMID:35305296
- 476 16. Zhang J, Chen N, Zhao D, Zhang J, Hu Z, Tao Z. Clinical Characteristics of COVID-19 Patients
477 Infected by the Omicron Variant of SARS-CoV-2. *Front Med* 2022 May 9;9:912367.
478 PMID:35615088
- 479 17. Whitaker M, Elliott J, Bodinier B, Barclay W, Ward H, Cooke G, Donnelly CA, Chadeau-Hyam M,
480 Elliott P. Variant-specific symptoms of COVID-19 in a study of 1,542,510 adults in England. *Nat*
481 *Commun* 2022 Nov 11;13(1):6856. PMID:36369151
- 482 18. Hitchings MDT, Ranzani OT, Torres MSS, de Oliveira SB, Almiron M, Said R, Borg R, Schulz
483 WL, de Oliveira RD, da Silva PV, de Castro DB, Sampaio V de S, de Albuquerque BC, Ramos
484 TCA, Fraxe SHH, da Costa CF, Naveca FG, Siqueira AM, de Araújo WN, Andrews JR, Cummings
485 DAT, Ko AI, Croda J. Effectiveness of CoronaVac among healthcare workers in the setting of high
486 SARS-CoV-2 Gamma variant transmission in Manaus, Brazil: A test-negative case-control study.
487 *Lancet Reg Health Am* 2021 Sep;1:100025. PMID:34386791
- 488 19. Vahey GM, Marshall KE, McDonald E, Martin SW, Tate JE, Midgley CM, Killerby ME, Kawasaki
489 B, Herlihy RK, Alden NB, Staples JE, Colorado Investigation Team2. Symptom Profiles and
490 Progression in Hospitalized and Nonhospitalized Patients with Coronavirus Disease, Colorado,
491 USA, 2020. *Emerg Infect Dis* 2021 Feb;27(2):385–395. PMID:33496225
- 492 20. Vihta KD, Pouwels KB, Peto TE, Pritchard E, House T, Studley R, Rourke E, Cook D, Diamond I,
493 Crook D, Clifton DA, Matthews PC, Stoesser N, Eyre DW, Walker AS, COVID-19 Infection
494 Survey team. Omicron-associated changes in SARS-CoV-2 symptoms in the United Kingdom. *Clin*
495 *Infect Dis* 2022 Aug 3;76(3):e133–41. PMID:35917440
- 496 21. Menni C, Valdes AM, Polidori L, Antonelli M, Penamakuri S, Nogal A, Louca P, May A,
497 Figueiredo JC, Hu C, Molteni E, Canas L, Österdahl MF, Modat M, Sudre CH, Fox B, Hammers A,
498 Wolf J, Capdevila J, Chan AT, David SP, Steves CJ, Ourselin S, Spector TD. Symptom prevalence,
499 duration, and risk of hospital admission in individuals infected with SARS-CoV-2 during periods of
500 omicron and delta variant dominance: a prospective observational study from the ZOE COVID
501 Study. *Lancet* 2022 Apr 23;399(10335):1618–1624. PMID:35397851
- 502 22. Malden DE, Tartof SY, Ackerson BK, Hong V, Skarbinski J, Yau V, Qian L, Fischer H, Shaw SF,
503 Caparosa S, Xie F. Natural Language Processing for Improved Characterization of COVID-19
504 Symptoms: Observational Study of 350,000 Patients in a Large Integrated Health Care System.
505 *JMIR Public Health Surveill* 2022 Dec 30;8(12):e41529. PMID:36446133
- 506 23. World Health Organization. Statement on the fifteenth meeting of the IHR (2005) Emergency
507 Committee on the COVID-19 pandemic [Internet]. 2023 [cited 2023 May 14]. Available from:
508 [https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-](https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-)
509 [health-regulations-\(2005\)-emergency-committee-regarding-the-coronavirus-disease-\(covid-19\)-](https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-)

- 510 pandemic
- 511 24. Meads S, McLemore T. The national ambulatory medical care survey: symptom classification. *Vital*
512 *Health Stat 2* 1974 May;(63):1–41. PMID:25093394
- 513 25. Rennert L, Ma Z, McMahan CS, Dean D. Effectiveness and protection duration of Covid-19
514 vaccines and previous infection against any SARS-CoV-2 infection in young adults. *Nat Commun*
515 2022 Jul 8;13(1):3946. PMID:35803915
- 516 26. Friedman JH, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via
517 Coordinate Descent. *J Stat Softw* 2010 Feb 2;33:1–22.
- 518 27. Looi M-K. How are covid-19 symptoms changing? *BMJ* 2023 Jan 18;380:3. PMID:36653024
- 519 28. Marquez C, Kerkhoff AD, Schrom J, Rojas S, Black D, Mitchell A, Wang C-Y, Pilarowski G,
520 Ribeiro S, Jones D, Payan J, Manganelli S, Rojas S, Lemus J, Jain V, Chamie G, Tulier-Laiwa V,
521 Petersen M, DeRisi J, Havlir DV. COVID-19 Symptoms and Duration of Rapid Antigen Test
522 Positivity at a Community Testing and Surveillance Site During Pre-Delta, Delta, and Omicron
523 BA.1 Periods. *JAMA Netw Open* 2022 Oct 3;5(10):e2235844. PMID:36215069
- 524 29. Callejon-Leblic MA, Moreno-Luna R, Del Cuvillo A, Reyes-Tejero IM, Garcia-Villaran MA,
525 Santos-Peña M, Maza-Solano JM, Martín-Jimenez DI, Palacios-Garcia JM, Fernandez-Velez C,
526 Gonzalez-Garcia J, Sanchez-Calvo JM, Solanellas-Soler J, Sanchez-Gomez S. Loss of Smell and
527 Taste Can Accurately Predict COVID-19 Infection: A Machine-Learning Approach. *J Clin Med Res*
528 2021 Feb 3;10(4). PMID:33546319
- 529 30. Struyf T, Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MM, Spijker R, Hooft L,
530 Emperador D, Dittrich S, Domen J, Horn SRA, Van den Bruel A, Cochrane COVID-19 Diagnostic
531 Test Accuracy Group. Signs and symptoms to determine if a patient presenting in primary care or
532 hospital outpatient settings has COVID-19 disease. *Cochrane Database Syst Rev* 2020 Jul
533 7;7(7):CD013665. PMID:32633856
- 534 31. Ellen Fragaszy, Madhumita Shrotri, Cyril Geismar, Anna Aryee, Sarah Beale, Isobel Braithwaite,
535 Thomas Byrne, Max T. Eyre, Wing Lam Erica Fong, Jo Gibbs, Pia Hardelid, Jana Kovar, Vasileios
536 Lampos, Eleni Nastouli, Annalan M.D. Navaratnam, Vincent Nguyen, Parth Patel, Robert W.
537 Aldridge, Andrew Hayward, Virus Watch Collaborative. Symptom profiles and accuracy of clinical
538 case definitions for COVID-19 in a community cohort: results from the Virus Watch study [version
539 1; peer review: awaiting peer review]. 2022. Available from:
540 <https://wellcomeopenresearch.org/articles/7-84> [accessed Apr 7, 2023]
- 541 32. Clarke KEN, Jones JM, Deng Y, Nycz E, Lee A, Iachan R, Gundlapalli AV, Hall AJ, MacNeil A.
542 Seroprevalence of Infection-Induced SARS-CoV-2 Antibodies - United States, September 2021-
543 February 2022. *MMWR Morb Mortal Wkly Rep* 2022 Apr 29;71(17):606–608. PMID:35482574
- 544 33. Tso WWY, Kwan MYW, Wang YL, Leung LK, Leung D, Chua GT, Ip P, Fong DYT, Wong WHS,
545 Chan SHS, Chan JFW, Peiris M, Lau YL, Rosa Duque JS. Severity of SARS-CoV-2 Omicron BA.2
546 infection in unvaccinated hospitalized children: comparison to influenza and parainfluenza
547 infections. *Emerg Microbes Infect* 2022 Dec;11(1):1742–1750. PMID:35730665
- 548 34. Krutikov M, Stirrup O, Nacer-Laidi H, Azmi B, Fuller C, Tut G, Palmer T, Shrotri M, Irwin-Singer
549 A, Baynton V, Hayward A, Moss P, Copas A, Shallcross L, COVID-19 Genomics UK consortium.
550 Outcomes of SARS-CoV-2 omicron infection in residents of long-term care facilities in England
551 (VIVALDI): a prospective, cohort study. *Lancet Healthy Longev* 2022 May;3(5):e347–e355.
552 PMID:35531432
- 553 35. Mancilla-Galindo J, Kammar-García A, Martínez-Esteban A, Meza-Comparán HD, Mancilla-

- 554 Ramírez J, Galindo-Sevilla N. COVID-19 patients with increasing age experience differential time
555 to initial medical care and severity of symptoms. *Epidemiol Infect* 2021 Oct 22;149:e230.
556 PMID:34674789
- 557 36. Marin-Gomez FX, Mendioroz-Peña J, Mayer M-A, Méndez-Boo L, Mora N, Hermosilla E, Coma E,
558 Vilaseca J-M, Leis A, Medina M, Miró Catalina Q, Vidal-Alaball J. Comparing the Clinical
559 Characteristics and Mortality of Residential and Non-Residential Older People with COVID-19:
560 Retrospective Observational Study. *Int J Environ Res Public Health* 2022 Jan 2;19(1).
561 PMID:35010742
- 562 37. Tam EMY, Kwan YK, Ng YY, Yam PW. Clinical course and mortality in older patients with
563 COVID-19: a cluster-based study in Hong Kong. *Hong Kong Med J* 2022 Jun;28(3):215–222.
564 PMID:35686472
- 565 38. Patgiri PR, Rajendran V, Ahmed AB. Clinico-Epidemiological Profiles of COVID-19 Elderly
566 Patients in Guwahati City, Assam, India: A Cross-Sectional Study. *Cureus* 2022 Apr;14(4):e24043.
567 PMID:35573552
- 568 39. Almayahi ZK, Raveendran AV, Al Malki R, Safwat A, Al Baloshi M, Abbas A, Al Salami AS, Al
569 Mujaini SM, Al Dhuhli K, Al Mandhari S. Clinical features, laboratory characteristics and risk
570 factors for mortality of COVID-19 patients in a secondary hospital in Oman during the first wave of
571 the SARS-CoV-2 pandemic. *Bull Natl Salmon Resour Cent* 2022 May 16;46(1):139.
572 PMID:35601475
- 573 40. Ryan C, Minc A, Caceres J, Balsalobre A, Dixit A, Ng BK, Schmitzberger F, Syed-Abdul S, Fung
574 C. Predicting severe outcomes in Covid-19 related illness using only patient demographics,
575 comorbidities and symptoms. *Am J Emerg Med* 2021 Jul;45:378–384. PMID:33046294
- 576 41. Cuadros DF, Moreno CM, Musuka G, Miller FD, Coule P, MacKinnon NJ. Association Between
577 Vaccination Coverage Disparity and the Dynamics of the COVID-19 Delta and Omicron Waves in
578 the US. *Front Med* 2022 Jun 14;9:898101. PMID:35775002
- 579 42. Kwok KO, Cowling B, Wei V, Riley S, Read JM. Temporal variation of human encounters and the
580 number of locations in which they occur: a longitudinal study of Hong Kong residents. *J R Soc
581 Interface* 2018 Jan;15(138). PMID:29367241
- 582 43. Li J, Huang DQ, Zou B, Yang H, Hui WZ, Rui F, Yee NTS, Liu C, Nerurkar SN, Kai JCY, Teng
583 MLP, Li X, Zeng H, Borghi JA, Henry L, Cheung R, Nguyen MH. Epidemiology of COVID-19: A
584 systematic review and meta-analysis of clinical characteristics, risk factors, and outcomes. *J Med
585 Virol* 2021 Mar;93(3):1449–1458. PMID:32790106
- 586 44. Kwok KO, McNeil EB, Tsoi MTF, Wei VWI, Wong SYS, Tang JWT. Will achieving herd
587 immunity be a road to success to end the COVID-19 pandemic? *J Infect* 2021 Sep;83(3):381–412.
588 PMID:34118274

TABLES AND FIGURES

Table 1. Demographic and epidemiological characteristics of all included COVID-19 cases.

Characteristic	Overall N = 76,762	Wild type N = 1,175	Delta N = 284	Omicron (BA.1) N = 305	Omicron (BA.2) N = 74,685	Omicron (BA.4) N = 63	Omicron (BA.5) N = 250
Age group (years)							
0-9	4,397 (5.7)	33 (2.8)	12 (4.2)	9 (3.0)	4,332 (5.8)	0 (0.0)	11 (4.4)
10-19	3,649 (4.8)	172 (14.6)	17 (6.0)	19 (6.2)	3,415 (4.6)	4 (6.3)	22 (8.8)
20-29	11,785 (15.4)	280 (23.8)	35 (12.3)	77 (25.2)	11,337 (15.2)	14 (22.2)	42 (16.8)
30-39	17,046 (22.2)	249 (21.2)	29 (10.2)	63 (20.7)	16,645 (22.3)	3 (4.8)	57 (22.8)
40-49	15,015 (19.6)	134 (11.4)	44 (15.5)	71 (23.3)	14,704 (19.7)	13 (20.6)	49 (19.6)
50-59	13,148 (17.1)	146 (12.4)	33 (11.6)	38 (12.5)	12,876 (17.2)	13 (20.6)	42 (16.8)
60-69	7,463 (9.7)	109 (9.3)	46 (16.2)	20 (6.6)	7,263 (9.7)	10 (15.9)	15 (6.0)
70+	4,259 (5.5)	52 (4.4)	68 (23.9)	8 (2.6)	4,113 (5.5)	6 (9.5)	12 (4.8)
Sex							
Female	40,968 (53.4)	543 (46.2)	154 (54.2)	155 (50.8)	39,954 (53.5)	27 (42.9)	135 (54.0)
Male	35,794 (46.6)	632 (53.8)	130 (45.8)	150 (49.2)	34,731 (46.5)	36 (57.1)	115 (46.0)
Wave							
1-2	1,175 (1.5)	1,175 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
5	75,587 (98.5)	0 (0.0)	284 (100.0)	305 (100.0)	74,685 (100.0)	63 (100.0)	250 (100.0)
Classification ^a							
Imported ^b	3,265 (4.3)	770 (65.5)	14 (4.9)	304 (99.7)	2,061 (2.8)	30 (47.6)	86 (34.4)
Local	73,497 (95.7)	405 (34.5)	270 (95.1)	1 (0.3)	72,624 (97.2)	33 (52.4)	164 (65.6)
Vaccination status ^c							
Unvaccinated	14,001 (18.2)	1,175 (100.0)	95 (33.5)	38 (12.5)	12,677 (17.0)	3 (4.8)	13 (5.2)
One complete dose	4,453 (5.8)	0 (0.0)	31 (10.9)	8 (2.6)	4,411 (5.9)	0 (0.0)	3 (1.2)
One incomplete dose	2,002 (2.6)	0 (0.0)	6 (2.1)	0 (0.0)	1,996 (2.7)	0 (0.0)	0 (0.0)
Two complete doses	38,425 (50.1)	0 (0.0)	130 (45.8)	183 (60.0)	38,057 (51.0)	11 (17.5)	44 (17.6)
One complete and one incomplete dose	1,473 (1.9)	0 (0.0)	1 (0.4)	0 (0.0)	1,472 (2.0)	0 (0.0)	0 (0.0)
Three complete doses	12,898 (16.8)	0 (0.0)	15 (5.3)	69 (22.6)	12,578 (16.8)	48 (76.2)	188 (75.2)
Two complete and one incomplete doses	3,510 (4.6)	0 (0.0)	6 (2.1)	7 (2.3)	3,494 (4.7)	1 (1.6)	2 (0.8)
Presence of symptoms							
No	22,312 (29.1)	302 (25.7)	72 (25.4)	131 (43.0)	21,690 (29.0)	32 (50.8)	85 (34.0)
Yes	54,450 (70.9)	873 (74.3)	212 (74.6)	174 (57.0)	52,995 (71.0)	31 (49.2)	165 (66.0)

^a Local cases consist of "local cases", "locally acquired cases", "possibly local cases", and "close contact of (possibly) local cases". Others were grouped as imported cases.

^b The majority of imported cases were from the United Kingdom (13.4%), the United States (7.7%), and Singapore (4.4%).

^c A complete dose refers to receipt of a vaccine dose given at least 14 days before the symptom onset date (for symptomatic cases) or the report date (for asymptomatic cases).

Table 2. Association of symptoms with death among unvaccinated symptomatic cases

Symptom	Frequency			Logistic regression		
	Overall (N=7,988)	Alive (N=7,913)	Dead (N=75)	OR	95% CI	
General symptoms						
Chills	258 (3.2%)	257 (3.2%)	1 (1.3%)	0.68	0.04	3.31
Fatigue	251 (3.1%)	249 (3.1%)	2 (2.7%)	0.58	0.09	2.03
Fever	5,080 (63.6%)	5,032 (63.6%)	48 (64.0%)	2.04	1.22	3.47
Malaise	41 (0.5%)	40 (0.5%)	1 (1.3%)	2.06	0.11	12.0
Nervous system						
Disturbance of taste and/or smell	104 (1.3%)	103 (1.3%)	1 (1.3%)	5.85	0.31	31.6
Dizziness	145 (1.8%)	144 (1.8%)	1 (1.3%)	0.78	0.04	4.07
Headache	758 (9.5%)	756 (9.6%)	2 (2.7%)	0.39	0.06	1.33
Respiratory system						
Blocked nose	139 (1.7%)	138 (1.7%)	1 (1.3%)	3.67	0.20	19.1
Chest pain	262 (3.3%)	261 (3.3%)	1 (1.3%)	0.43	0.02	2.13
Cough	3,350 (41.9%)	3,322 (42.0%)	28 (37.3%)	0.51	0.30	0.84
Runny nose	1,162 (14.5%)	1,157 (14.6%)	5 (6.7%)	0.53	0.18	1.25
Shortness of breath	538 (6.7%)	519 (6.6%)	19 (25.3%)	2.25	1.21	4.03
Sputum (expectoration)	417 (5.2%)	415 (5.2%)	2 (2.7%)	0.38	0.06	1.28
Digestive system						
Abdominal pain	66 (0.8%)	65 (0.8%)	1 (1.3%)	4.28	0.21	27.6
Nausea or vomiting	332 (4.2%)	327 (4.1%)	5 (6.7%)	2.04	0.64	5.31
Sore throat	2,232 (27.9%)	2,224 (28.1%)	8 (10.7%)	0.43	0.19	0.87
Musculoskeletal system						
Back pain	30 (0.4%)	29 (0.4%)	1 (1.3%)	14.2	0.73	86.0
Unspecified muscle pain	235 (2.9%)	234 (3.0%)	1 (1.3%)	0.71	0.04	3.42

Remarks: All models were adjusted for age, sex, subvariant, and reporting delay.

Table 3. Average penalized LASSO estimates for symptoms associated with death and stepwise generalized linear models among unvaccinated elderly cases

	All Elderly n=1,852		Omicron (BA.2) n=1,721	
	LASSO	GLM	LASSO	GLM
	Avg est.	OR (95% CI)	Avg est.	OR (95% CI)
Predictive symptoms				
Abdominal pain	1.44	6.52 (0.30-44.7)	-	-
Blocked nose	1.11	5.24 (0.78-19.9)	1.19	5.34 (0.80-20.1)
Malaise	0.94	3.31 (0.17-19.7)	-	-
Shortness of breath	0.60	2.12 (1.20-3.57)	0.53	2.02 (1.12-3.47)
Fever (temperature>38°C)	0.32	1.49 (0.99-2.21)	0.23	1.38 (0.90-2.06)
Protective symptoms				
Cough	-0.79	0.36 (0.21-0.57)	-0.84	0.33 (0.19-0.54)
Diarrhea	-0.73	-	-0.56	-
Chest pain	-0.68	0.28 (0.04-0.99)	-0.63	0.30 (0.05-1.04)
Voice disorders	-0.64	-	-0.72	-
Sore throat	-0.52	0.43 (0.19-0.86)	-0.48	0.45 (0.20-0.90)
Loss of appetite	-0.46	-	-0.20	-
AIC		1218.75		1172.90
AUC		0.9261		0.9264

Avg est.: Average estimated coefficient. GLM: generalised linear model. LASSO: least absolute shrinkage and selection operator.

Remarks: 1. All models were adjusted for age, sex, presence of chronic diseases, and reporting delay.

2. Dashes indicate that the symptom was excluded in the model due to convergence failure.

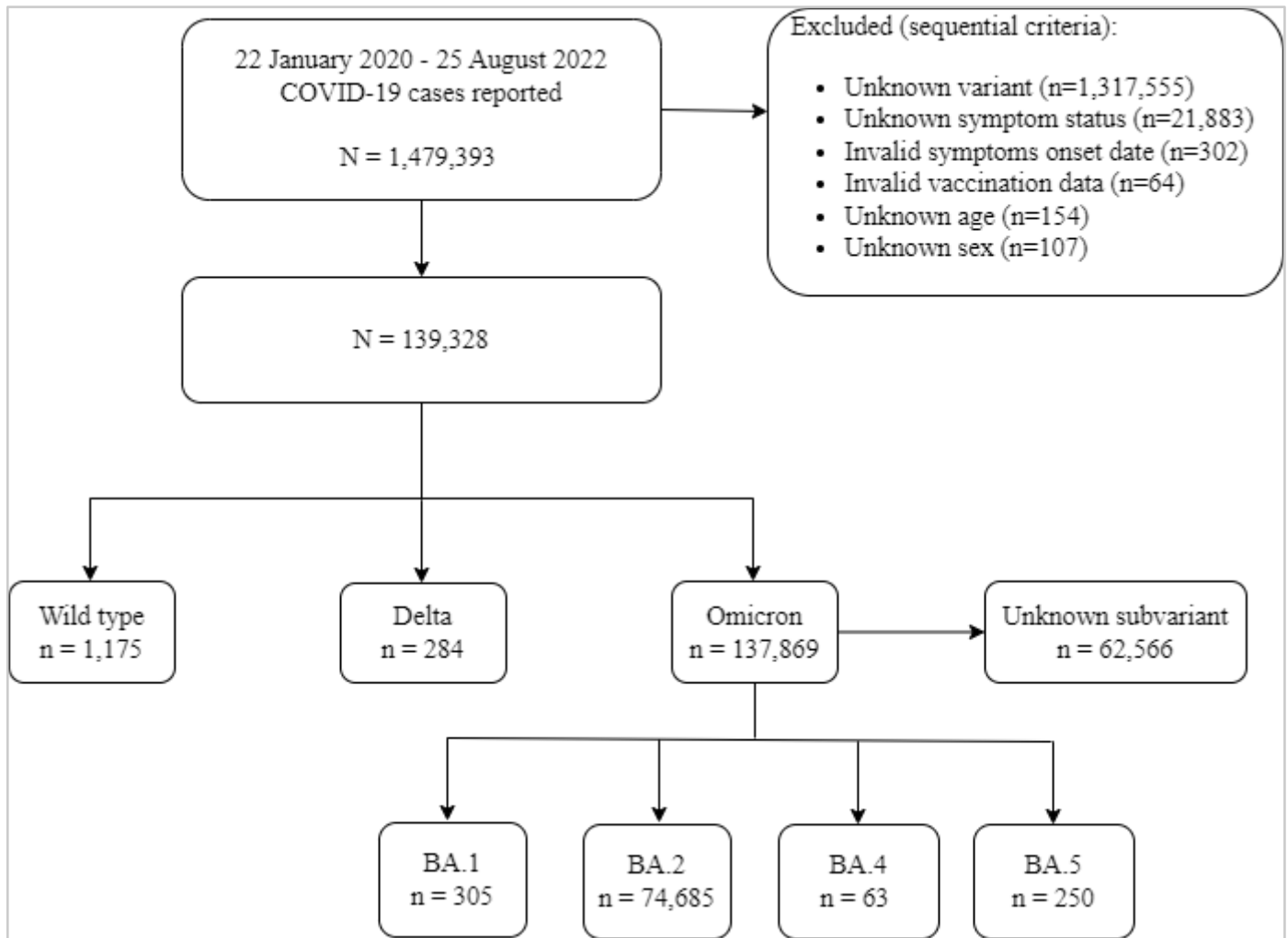


Figure 1. Study flowchart

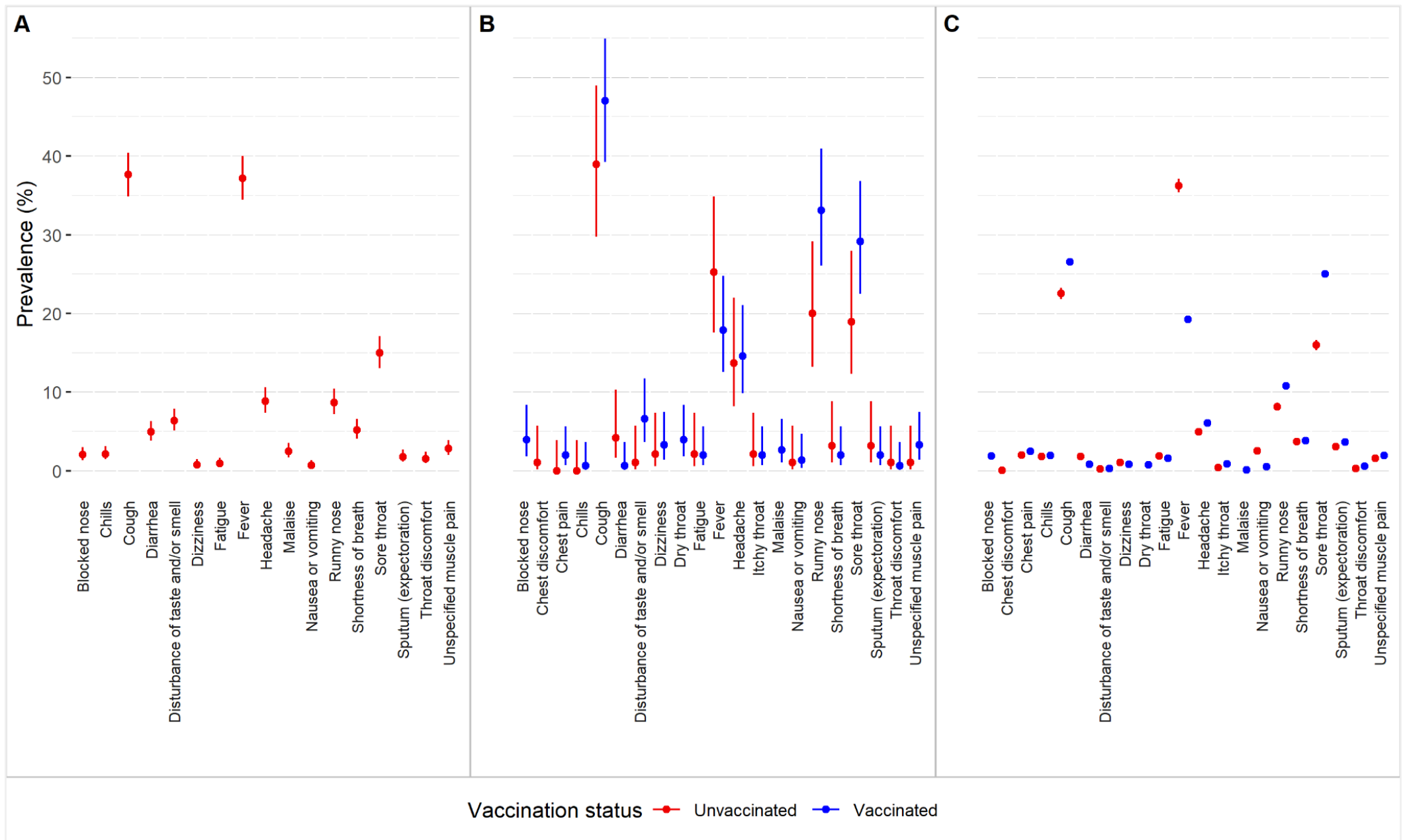


Figure 2. Prevalence of commonly reported symptoms among COVID-19 cases stratified by vaccination status for (A) wild-type, (B) delta and (C) omicron BA.2 variants. Only symptoms with a variant-specific combined (both the vaccinated and unvaccinated) prevalence >1% are shown.

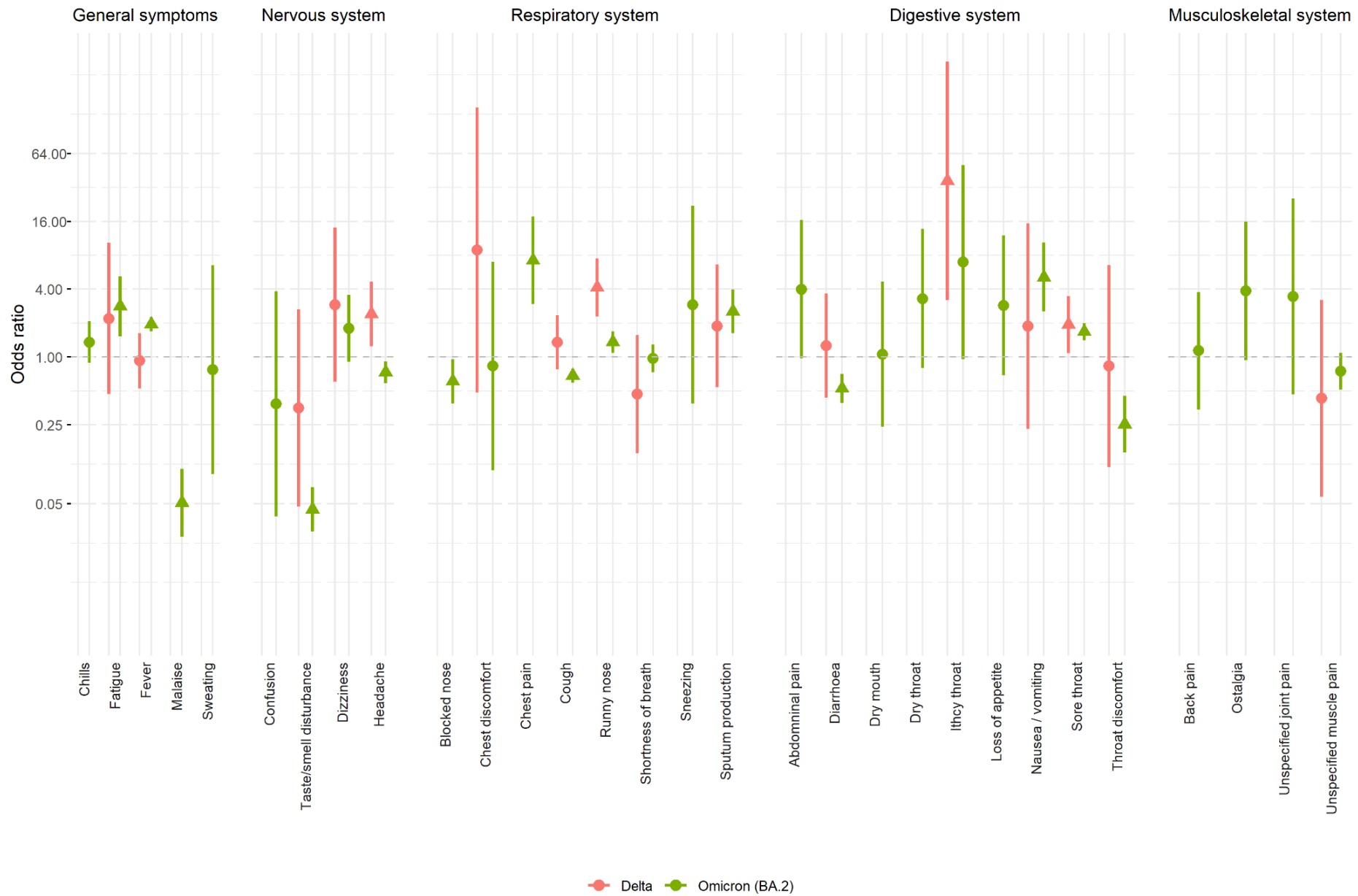


Figure 3. Effect of variant on symptom development among unvaccinated symptomatic cases (adjusting for age, sex, and reporting delay). Estimates that are significant and not significant are represented by solid triangles and circles, respectively.

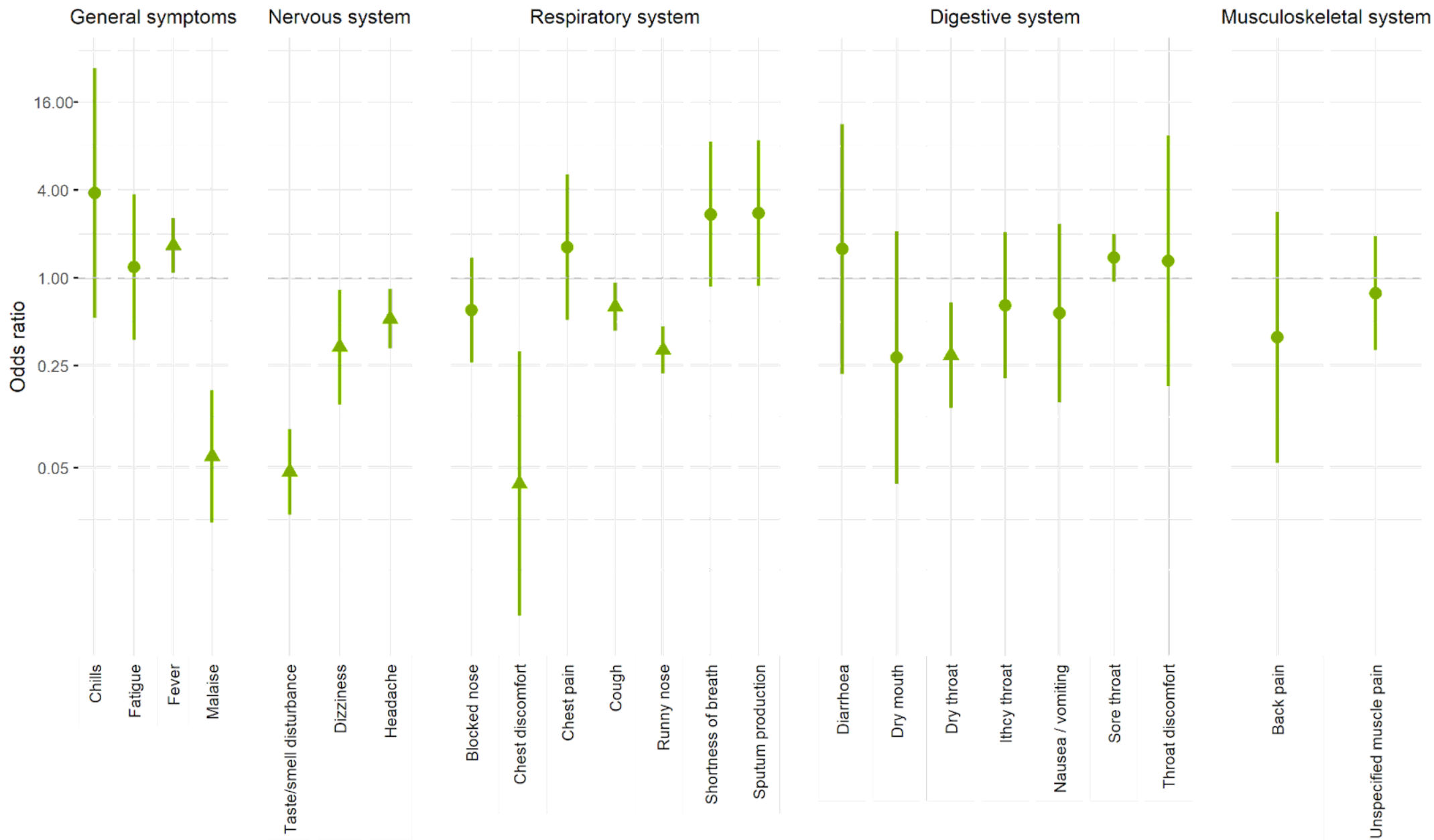


Figure 4. Effect of the omicron variant compared to delta on symptom development among fully vaccinated symptomatic cases. Solid triangles and circles represent estimates that are significant and not significant. (adjusting for age, sex, and reporting delay).

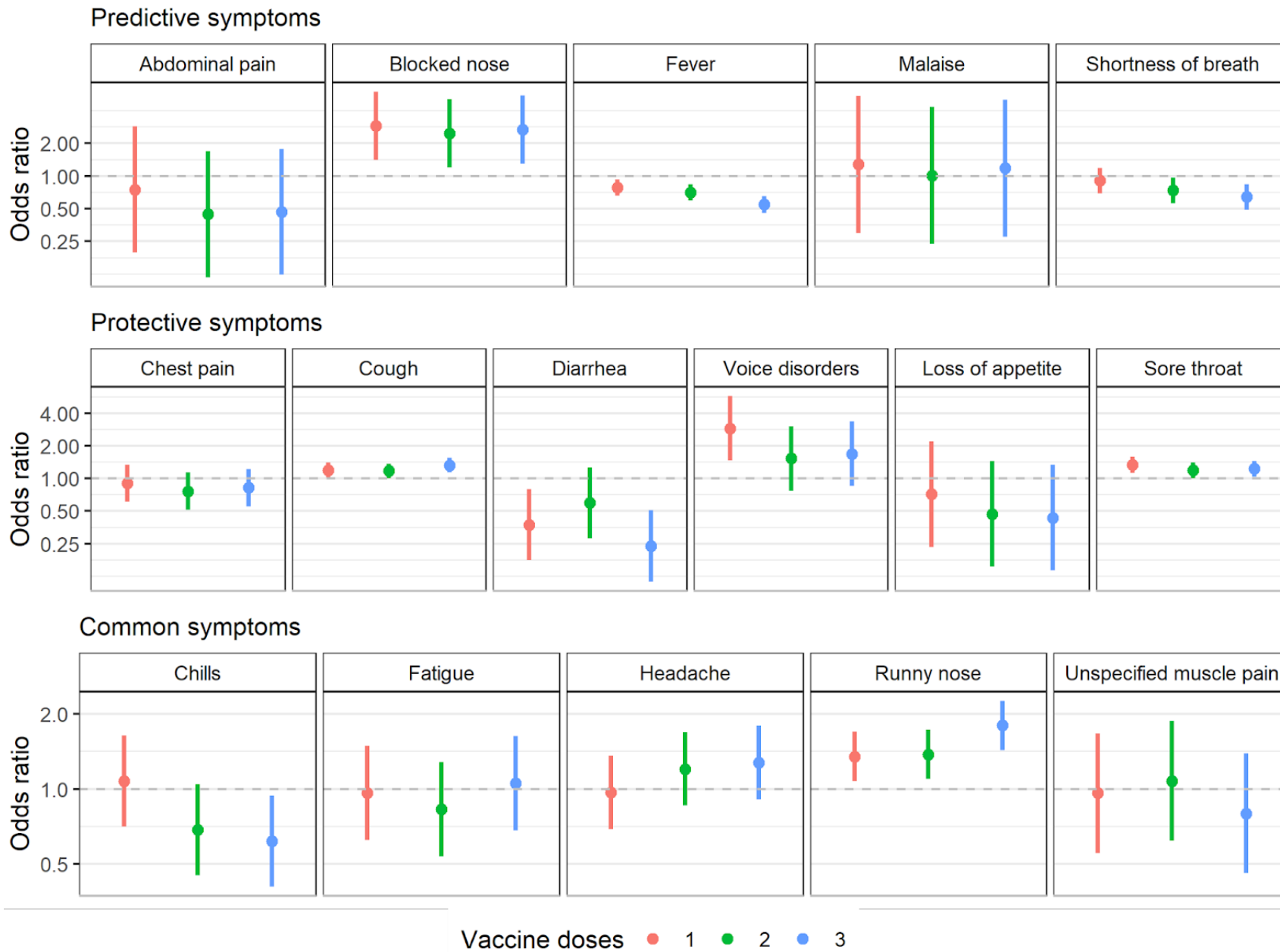


Figure 5. Effect of the number of vaccine doses on symptom development among symptomatic elderly cases, adjusting for age, sex, variant, and reporting delay. The reference group is the unvaccinated cases.

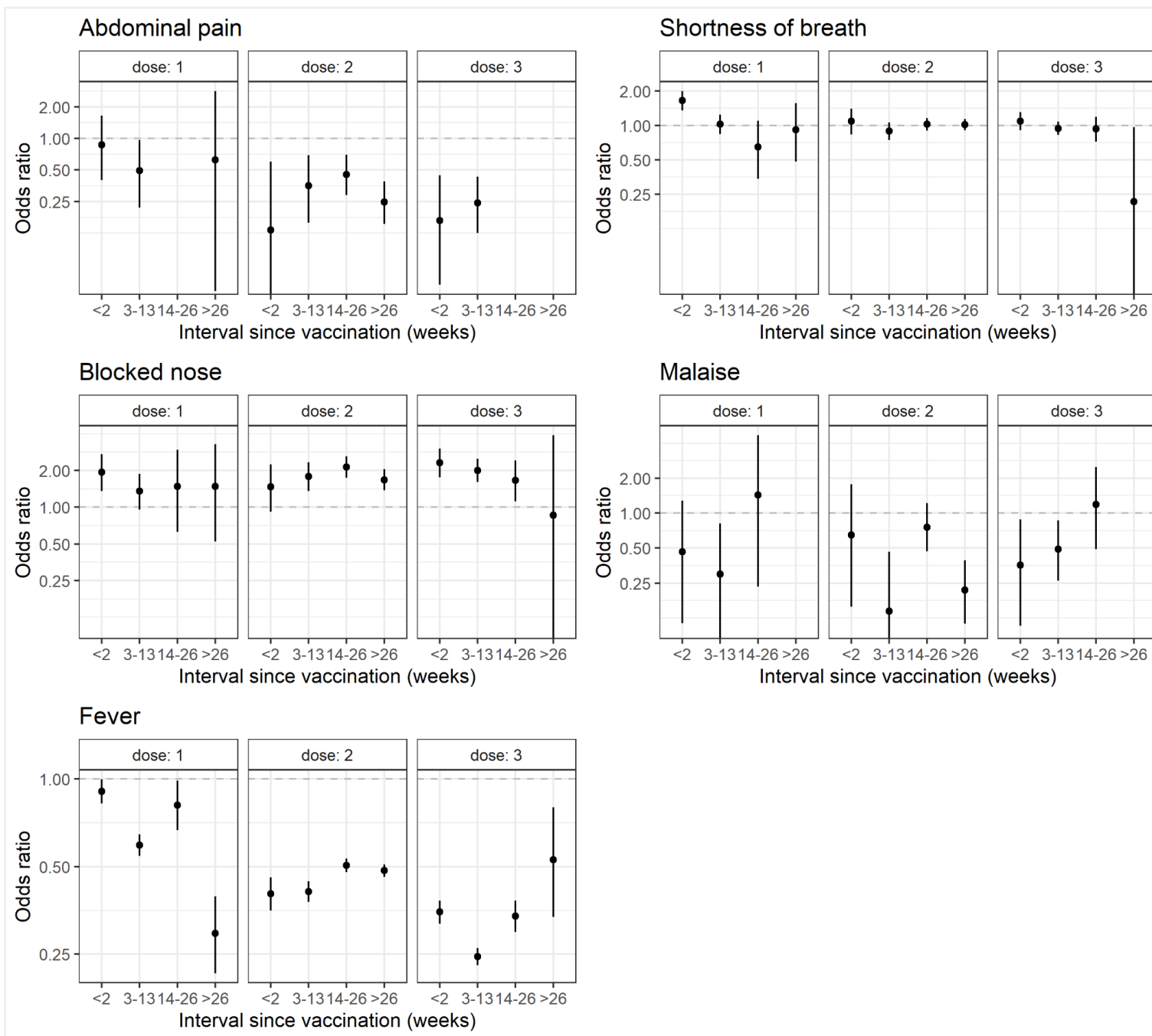


Figure 6. Effect of the time elapsed since the last vaccination on symptom development among symptomatic elderly cases, adjusting for age, sex, variant, and reporting delay. The reference group is the unvaccinated cases.