Impact of Pulmonary and Sleep Disorders on COVID-19 Infection Severity in a

Large Clinical Biobank

Brian E. Cade^{1,2*}, Syed Moin Hassan^{1,2,3}, Janet M. Mullington^{2,4}, Elizabeth W. Karlson^{5,6},

Susan Redline^{1,2}

- 1. Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Boston, MA 02115, USA
- 2. Division of Sleep Medicine, Harvard Medical School, Boston, MA 02115, USA
- 3. Division of Pulmonary Disease and Critical Care Medicine, University of Vermont, Burlington, VT 05401, USA
- 4. Department of Neurology, Beth Israel Deaconess Medical Center, Boston, MA 02115, USA
- 5. Division of Rheumatology, Inflammation and Immunity, Brigham and Women's Hospital, Boston, MA 02115, USA
- 6. Center for Genomic Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA

*: Please address all correspondence and requests for reprints to:

Brian Cade

Division of Sleep and Circadian Disorders

Brigham and Women's Hospital, Harvard Medical School

221 Longwood Avenue, Boston, MA 02115, USA

1-857-307-0353 (phone) 1-617-278-6946 (fax) bcade@bwh.harvard.edu (e-mail)

Author Contributions: Conception and design: BEC, SMH, EWK, and SR. Data

acquisition: all authors. Analysis: BEC. Interpretation, draft and review, and final approval:

all authors. BEC and SR had full access to the study data and take responsibility for the

integrity of the data and accuracy of analyses.

Funding: Brian Cade is supported by grants from the American Thoracic Society Foundation

and the National Institutes of Health (R03-HL154284, R01-HL153805). Susan Redline is supported by grants from the National Institutes of Health (R35-HL135818). Elizabeth Karlson is supported by grants from the National Institutes of Health (U01-HG008685 and P30 AR070253).

Running title: Common pulmonary disease and Covid-19 severity

Competing Interest Statement: The authors declare no competing interest.

Cade et al.: Common pulmonary disease and Covid-19 severity

1 Abstract

Rationale: Multiple pulmonary, sleep, and other disorders are associated with the
severity of Covid-19 infections but may or may not directly affect the etiology of acute
Covid-19 infection. Identifying the relative importance of concurrent risk factors may
prioritize respiratory disease outbreaks research.

6 **Objectives:** To identify associations of common preexisting pulmonary and sleep 7 disease on acute Covid-19 infection severity, investigate the relative contributions of each disease and selected risk factors, identify sex-specific effects, and examine whether 8 9 additional electronic health record (EHR) information would affect these associations. Methods: 45 pulmonary and 6 sleep diseases were examined in 37,020 patients with 10 11 Covid-19. We analyzed three outcomes: death; a composite measure of mechanical ventilation and/or ICU admission; and inpatient admission. The relative contribution of 12 pre-infection covariates including other diseases, laboratory tests, clinical procedures, 13 and clinical note terms was calculated using LASSO. Each pulmonary/sleep disease 14 model was then further adjusted for covariates. 15

Measurements and main results: 37 pulmonary/sleep diseases were associated
with at least one outcome at Bonferroni significance, 6 of which had increased relative
risk in LASSO analyses. Multiple prospectively collected non-pulmonary/sleep diseases,
EHR terms and laboratory results attenuated the associations between preexisting
disease and Covid-19 infection severity. Adjustment for counts of prior "blood urea
nitrogen" phrases in clinical notes attenuated the odds ratio point estimates of 12
pulmonary disease associations with death in women by ≥1.

23 Conclusions: Pulmonary diseases are commonly associated with Covid-19 infection

Cade et al.: Common pulmonary disease and Covid-19 severity

- 24 severity. Associations are partially attenuated by prospectively-collected EHR data,
- 25 which may aid in risk stratification and physiological studies.

26

- 27 Abstract word count: 250
- 28 Keywords: COVID-19, Epidemiology, Pulmonary disorders, Sleep disorders

29

30

Cade et al.: Common pulmonary disease and Covid-19 severity

31 Introduction

Covid-19 continues to be an important public health problem. Prior to the Omicron wave, Covid-19 was one of the leading causes of death in the US in every age group above age 15 [1]. Thousands continue to die despite the availability of vaccines, emphasizing a continued need for research to help mitigate the effects of this and future infectious disease outbreaks.

The range of effects arising from Covid-19 infection, from minimal symptoms to 37 38 death is striking. Less is known about the relative importance of individual pulmonary and sleep diseases contributing to infection severity in patients with multimorbidity. 39 Examining the association of a range of pulmonary and sleep diseases with Covid-19 40 41 infection severity and applying machine learning methods to identify a subset of 42 independent associations that predict infection severity may aid in identifying 43 etiological mechanisms of morbidity and mortality. These associations may differ by sex, 44 given sex differences in both Covid-19 prevalence and associated risk factors.

Associations between preexisting pulmonary/sleep disease and Covid-19 45 46 infection severity may be changing with time. The Omicron variant has reduced pulmonary pathology compared to other variants [2,3]. Patients carrying the Omicron 47 48 variant have reduced risk of hospital admissions and severe infection [4]. The degree of 49 association between individual diseases and Covid-19 severity may therefore change across different phases of the epidemic. Examining these changing environmental 50 interactions may provide insight into the pathophysiology of both Covid-19 and the 51 tested diseases themselves. 52

Cade et al.: Common pulmonary disease and Covid-19 severity

53	Prospective collection of data in clinical biobanks are one means of addressing
54	these and other questions. EHR databases can leverage extensive clinical note,
55	diagnosis, procedure, and laboratory data collected prior to infection. Natural language
56	processing (NLP) [5] of controlled medical vocabulary concepts [6] from date-stamped
57	clinical notes can improve disease phenotyping [7]. These clinical note concepts,
58	together with procedure and laboratory information, can also be used to identify risk
59	factors that complement or potentially mediate or confound the association of existing
60	diseases with Covid-19 infection severity.
61	In this study, we report associations between common preexisting pulmonary
62	and sleep diseases and acute Covid-19 morbidity and mortality among patients with
63	PCR-based diagnoses of Covid-19 with data in the Mass General Brigham (MGB) EHR
64	database from the greater Boston, MA region. We investigate the relative importance of
65	individual preexisting diseases and ancillary EHR concepts, identify sex- and date-
66	specific associations, and report the effect of lead EHR concepts on the associations of
67	preexisting diseases and Covid-19 infection severity.
68	
69	
70	Methods
71	Additional details are provided in the Supplemental Methods.
72	

73 <u>Study sample</u>

Cade et al.: Common pulmonary disease and Covid-19 severity All patients with PCR-diagnosed Covid-10 were enrolled in the MGB Research

/4	All patients with PCR-diagnosed Covid-19 were enrolled in the MGB Research
75	Patient Data Registry EHR database [8,9]. Study collection ended at 6/1/2022. We used
76	a stringent combined data floor and loyalty cohort [10] approach to minimize the
77	number of patients with reduced documentation (and increased risk of false negative
78	associations) [11] by requiring at least three clinical notes, encounters, and diagnoses.
79	We also required at least 7 clinical facts: one diagnosis, encounter, laboratory result,
80	medication, outpatient visit, routine visit [10], and vital sign obtained before 2020.
81	Children, MGB employees, and patients living outside of Massachusetts or without
82	available race/ethnicity, age, biological sex, or BMI values were excluded.
83	
84	Data used
85	The acute phase of infection for measuring ICU admissions, mechanical
86	ventilation, and inpatient admissions was defined as -7 to +30 days relative to the first
87	positive PCR test (we did not analyze Covid-19 reinfections) in the EHR. Diseases,
88	procedures, and NLP terms recorded at least 8 days prior to the PCR test were
89	extracted. Laboratory measurements were extracted from -5 to -1 years prior to the PCR
90	test. A breakthrough infection was defined as having occurred over 30 days after a
91	second vaccination. Majority variant date ranges were based on Broad Institute Covid-
92	19 Genome Surveillance project sequencing in the greater Boston area. Transition dates
93	from the pre-Delta to the Delta and Omicron variants were set at $7/3/2021$ and
94	12/17/2021 respectively.

95

- 4

Diseases were based on PheCode groupings of ICD codes [12,13] (Table S1) and

Cade et al.: Common pulmonary disease and Covid-19 severity

98	refined using natural language processing. We evaluated 45 common (\geq 1% prevalence in
99	patients with Covid-19) and selected rare pulmonary disorders and 6 common sleep
100	disorders [14–16]. Clinical note concept terms were extracted using NLP and mapped to
101	Concept Unique Identifiers (CUIs) codes [5,6]. Terms that matched disease names were
102	used to improve disease phenotyping using multimodal automated phenotyping (MAP)
103	[7]. We included all non-pulmonary and non-sleep diseases defined using PheCodes and
104	MAP with >1% prevalence among patients with Covid-19 in LASSO regression analyses
105	(described below). Given the sparse nature of certain individual procedures, we also
106	grouped procedures as specified by AHRQ Clinical Classification Software (CCS,
107	https://www.hcup-us.ahrq.gov/toolssoftware/ccs10/ccs10.jsp). Procedure and CUI IDs
108	for variables retained after LASSO regressions are listed in Table S6.
109	We documented evidence of continuous positive airway pressure (CPAP) use
110	among patients with at least one clinical encounter and one clinical note from -365 to -8
111	days prior to the first positive PCR test. Total counts of five related procedures and one
112	NLP term (Table S5) were combined. CPAP evidence in the prior year was present for
113	2,946 / 4,027 (73.1%) of the eligible patients with sleep apnea.
114	In an exploratory aim, we examined whether selected hematology laboratory
115	results were associated with Covid-19 severity. We included kidney function measures

given the retention of kidney disease in the LASSO results. The median values of
laboratory results for individual patients from -5 to -1 years were tested with and

118 without rank normalization. The 23 laboratory measurements were available for 65 -

119 89% of the patients (Table S8).

Cade et al.: Common pulmonary disease and Covid-19 severity

120

121 Statistical analyses

Logistic regression included 45 pulmonary disorders, 6 sleep disorders, was 122 adjusted for age, sex at birth, WHO obesity classification, race/ethnicity, and 123 breakthrough infection status to predict 3 outcomes: death, mechanical ventilation or 124 125 ICU admission, or inpatient admissions. The effect of CPAP evidence was tested using logistic regression with further adjustment for a single CPAP evidence term. 126 Adaptive LASSO regression used training and testing sets, each with cross-127 128 validation. Non-zero coefficient terms were carried forward to the training run and then 129 to the final combined sample run. Age, sex, BMI, race/ethnicity, breakthrough infection 130 status, healthcare encounter counts, and days into the pandemic were forced into runs. 131 Laboratory results were not included due to reduced sample size and non-random 132 ascertainment. We opted to include sex-specific terms in combined sex analyses and evaluate them in separate sex-stratified analyses. Finally, the potential effect of lead 133 LASSO variables on disease associations with Covid-19 infection severity was assessed 134 135 using logistic regression with standard covariates that included and excluded individual LASSO variables as covariates. 136

137

138

139 Results

140 <u>Sample characteristics</u>

Sample characteristics are listed in **Table 1**. The initial sample of patients with
Covid-19 (through 6/1/2022) and suitable data to clear a data floor was 37,020. Women

Cade et al.: Common pulmonary disease and Covid-19 severity

143 represented a majority of infections (62%) but a minority of patients with more severe 144 death and mechanical ventilation, or ICU admission outcomes (48%). The majority of breakthrough infections were in women (62%). At the time of the data freeze 145 (6/1/2022), the number of patients with an incident diagnosis during the pre-Delta 146 majority variant virus dates was slightly less than the combined number of patients with 147 148 a diagnosis during the Delta or Omicron majority variant diagnosis dates. The median BMI has decreased for Omicron majority variant diagnosis dates, as has the percentage 149 of Hispanic/Latino Americans among patients with an incident Covid-19 infection. The 150 151 prevalence of 45 pulmonary and 6 sleep diseases grouped by PheCodes (Table S1) may be reduced due to the strict case criteria used by NLP-defined phenotyping. The median 152 153 number of preexisting pulmonary and sleep diseases diagnosed for patients with an adverse Covid-19 outcome was higher than the median number of preexisting diseases 154 155 for all patients with Covid-19.

156

157 Associations between existing disease with Covid-19 infection severity

158 The associations between existing diseases and Covid-19 infection severity are described in **Table 2**. Models included adjustment for breakthrough infection, age, sex 159 160 at birth, obesity severity, and self-reported race/ethnicity. 42 of 51 diseases were nominally associated with any outcome (p < 0.05). 37 diseases were associated with an 161 162 outcome at Bonferroni significance ($p < 3.3 \times 10^{-4}$): 25, 23, and 37 diseases were associated with risk of death, mechanical ventilation or ICU admission, and inpatient 163 admission, respectively. Chronic pharyngitis and nasopharyngitis and septal 164 deviations / turbinate hypertrophy were inversely associated with risk of mechanical 165

Cade *et al.***:** Common pulmonary disease and Covid-19 severity ventilation or ICU admission at modest significance.

166

167 We analyzed sex-specific associations between existing diseases and Covid-19 infection severity. Death outcome results are shown in **Table 3**, with full results 168 provided in Table S2. 27 and 33 diseases were associated with at least one infection 169 severity outcome in men and women respectively at Bonferroni significance ($p < 1.1 \times$ 170 171 10⁻⁴). While multiple diseases displayed notably different odds ratio point estimates, sex 172 × disease interaction p-values were generally modest in this cohort with stringent loyalty and data floor criteria. Associations between existing diseases and Covid-19 173 174 infection severity in different age groups are provided in **Table S3**.

175 We examined associations between existing disease and Covid-19 incident 176 infection severity during days when the pre-Delta, Delta, and Omicron virus variants comprised the majority of samples sequenced from the greater Boston area. Table 4 177 178 lists the composite outcome results, with full results and sex-stratified results listed in Table S4. The pre-Delta, Delta, and Omicron majority variant date analyses were all 179 adjusted for breakthrough infection status. 31 and 26 diseases were associated with any 180 181 outcome during pre-Delta and Omicron variant majority dates respectively (Bonferroni $p < 1.1 \times 10^{-4}$). Seven diseases significantly associated with a death outcome during pre-182 Delta variant majority dates ($p < 1.1 \times 10^{-4}$) were notably attenuated (p > 0.01) during 183 Omicron variant majority dates (bacterial infection NOS, empyema and pneumothorax, 184 185 epistaxis or throat hemorrhage, pneumonia, pulmonary hypertension, sleep apnea, and symptoms involving respiratory system and other chest symptoms). Conversely, a 186 modest association between solitary pulmonary nodule and the death outcome during 187 pre-Delta variant majority dates strengthened to Bonferroni significance during 188

Cade et al.: Common pulmonary disease and Covid-19 severity

- 189 Omicron variant majority dates.
- 190

191 Effects of continuous airway positive pressure

We examined whether adjusting for evidence of continuous positive airway pressure (CPAP) usage in clinical notes and procedures in the year prior to infection would attenuate the associations with sleep apnea, given the trends we identified in our previous report with a smaller sample size [14]. All associations with sleep apnea were attenuated (p > 0.02, **Table S5**).

197

198 <u>Relative importance of individual existing diseases and EHR information</u>

199 Given the large number of associated diseases and elevated multimorbidity (18.8% of the patients carried ≥ 5 of the tested diseases), we analyzed the relative 200 201 importance of individual existing diseases affecting the risk of Covid-19 infection severity. We performed adaptive LASSO in the combined sex and sex-stratified datasets 202 by including all pulmonary and sleep diseases and multiple covariates. Non-zero 203 204 coefficients from the outcome results are shown in Tables 5 and S6, with larger 205 absolute values indicating greater importance for individual variables. Chronic airway obstruction, pneumonia, and respiratory failure had non-zero coefficients for the death 206 207 outcome.

We next leveraged additional EHR information by performing similar adaptive LASSO analyses to identify factors with greater importance using separate analyses of 505 non-pulmonary/sleep diseases, 1,647 procedures, 171 CCS grouped procedures, and 9,839 clinical note NLP terms. Results for non-zero coefficient terms are listed in **Table**

Cade et al.: Common pulmonary disease and Covid-19 severity

212	S6 . 124 EHR variables were identified in the combined sex analyses. Six non-
213	pulmonary/sleep diseases and ICD-coded procedures had non-zero coefficient terms in
214	the combined-sex death outcome analysis: chronic renal failure, chemotherapy, altered
215	mental status, essential hypertension, thrombocytopenia, and abnormality of gait.
216	Associations between LASSO terms and Covid-19 infection severity adjusted for
217	covariates are presented in Table S7. Multiple non-pulmonary diseases were associated
218	with the death outcome, including chronic renal failure (OR 3.61, 95% CI 3.04 – 4.28)
219	and thrombocytopenia (OR 3.19, 95% CI 2.45 – 4.17).
220	
221	Predictive laboratory values

222 Complete blood count and other hematology measurements are commonly collected in many patients undergoing hospital procedures. As an exploratory aim, we 223 224 examined whether the median values of these measurements collected 1 - 5 years prior to the incident Covid-19 diagnosis were associated with infection severity. We also 225 included measures of kidney function, given the identification of kidney disease in the 226 227 LASSO results. 17 of the 23 laboratory tests were associated with at least one outcome in the available sample at Bonferroni significance (65 – 89% coverage of patients), using 228 logistic regression with adjustment for covariates (Table S8). Estimated glomerular 229 230 filtration rate, hematocrit, hemoglobin, and red cell distribution width results were the 231 most notable associated measures with all p-values $\leq 2.0 \times 10^{-16}$, despite the identification of thrombocytopenia and elevated white blood cell counts PheCodes in the 232 233 LASSO results.

234

Cade *et al.*: Common pulmonary disease and Covid-19 severity 235 <u>Associated EHR concepts attenuate disease associations</u>

236 Finally, we examined whether individual retained LASSO terms and prior hematology and kidney laboratory results would attenuate the associations between 237 individual diseases and infection severity when included individually as model covariate 238 terms. The lead terms with the highest median point estimate adjustment differences 239 240 across all disease associations with the death outcome are shown in Figure 1. Lead terms spanned a range of healthcare procedures, laboratory results, and pulmonary and 241 non-pulmonary diseases. Full results are provided in Table S9, and additional 242 243 visualizations are provided in **Figures S1 – S17**. Adjustment for "blood urea nitrogen" 244 clinical note phases attenuated the odds ratio point estimates of 12 pulmonary disease 245 associations with death in women by \geq 1.0.

246

247

248 Discussion

In this study, we examined the association of preexisting pulmonary, sleep, and 249 250 other diseases with the risk of acute Covid-19 infection severity and investigated other factors that may alter those associations. Notable results include (a) common 251 associations between dozens of pulmonary disorders and infection severity, some of 252 which are protective; (b) changes in disease associations across date periods with 253 254 different SARS-CoV-2 majority variants, with attenuated associations with death for 7 diseases, and an increased point estimate for 1 disease; (c) an increased importance of 255 specific pulmonary disease associations with Covid-19 severity relative to other 256 comorbid pulmonary and sleep diseases; (d) associations between infection severity and 257

Cade *et al.*: Common pulmonary disease and Covid-19 severity

258	multiple hematology and kidney-related laboratory results collected $1-5$ years prior to
259	infection; and (e) non-pulmonary/sleep diseases, laboratory results, hospital
260	procedures, and clinical note terms that attenuate and sometimes increase the
261	associations between pulmonary and sleep diseases with Covid-19 infection severity.
262	These results may aid in further understanding the etiology of a changing disease.
263	We observed common associations between preexisting pulmonary and sleep
264	diseases with Covid-19 infection severity, including 37 of the 51 tested diseases
265	associated with at least one outcome at Bonferroni significance (Tables 2 and S2).
266	Certain diseases (e.g. asthma) can confer different risks for men and women, although
267	sex \times disease interaction p-values tended to be modest (Tables 3 and S2). Acute sinusitis
268	has a protective association with the death measure in men (0.33 OR, 95% CI 0.14 –
269	0.77, p = 9.8×10^{-3}) that is not present in women (0.94 OR, 95% CI 0.61 – 1.43, p =
270	0.76). A similar but more modest association was also observed between the death
271	outcome and acute upper respiratory infections. These findings may help improve the
272	personalization of treatment for patients with increased Covid-19 morbidity.
273	The associations between preexisting disease and infection severity are changing
274	across time for certain diseases (Tables 4 and S4). Notable diseases with attenuated
275	associations with death during Omicron-majority variant periods compared to pre-Delta
276	majority variant periods include bacterial infection NOS (3.00 OR [95% CI 2.16 - 4.16]

277 pre-Delta period; 1.63 [0.90 - 2.95] Omicron period), sleep apnea (1.83 [1.44 - 2.33] pre-

278 Delta period; 1.35 [0.89 - 2.04] Omicron period), and influenza (1.72 [1.29 - 2.30] pre-

279 Delta period; 0.90 [0.53 - 1.55] Omicron period). These changes may reflect differences

280 in SARS-CoV-2 pathogenicity in upper and lower respiratory tracts [2,3]. The

Page 15 of 40

Cade et al.: Common pulmonary disease and Covid-19 severity

differences in outcomes across majority variant strain periods are consistent with a recent report examining the same dataset using independent methods [17]. Differences in risk factors and outcomes have potential study design implications for acute and possibly Long Covid research. The changing nature of SARS-CoV-2 associations with a range of comorbidities could potentially provide a window for future physiological research examining how individual pulmonary and sleep diseases increase risk for outcomes in other diseases.

288 Our study sample has a high degree of multimorbidity, and 39 pre-existing diseases were associated with at least one infection severity outcome at Bonferroni 289 significance. We therefore performed LASSO regression analyses to identify the relative 290 291 importance of specific diseases (Tables 5 and S6). LASSO analyses minimize the number 292 of factors that collectively may predict an outcome. Six pulmonary diseases were 293 retained: chronic airway obstruction, respiratory failure, pneumonia, emphysema, 294 pulmonary congestion and hypostasis, and shortness of breath (Table 5). These diseases may be enriched with pathophysiological mechanisms that contribute increased risk 295 across the broader range of pulmonary and sleep disorders associated with acute Covid-296 19 infection severity. 297

Preexisting non-pulmonary diseases, hospital procedures and clinical note terms
identified by LASSO regression and laboratory results may have predictive value for
Covid-19 infection severity, given the common associations between preexisting
pulmonary disease and Covid-19 infection severity (Table S6). These variables may
potentially mediate or confound associations between preexisting pulmonary and sleep
diseases and infection severity (Table S9 and Figures 1, S1 – S17) or could reflect the

Cade et al.: Common pulmonary disease and Covid-19 severity

304 effects of other factors unmeasured in this study. Five non-pulmonary PheCode-305 encoded diseases (and chemotherapy) were associated with death: chronic renal failure, 306 altered mental status, essential hypertension, thrombocytopenia, and abnormality of gait. We chose to analyze procedures and all NLP term with common frequencies in 307 order to identify factors that could help predict outcome severity, even if they are 308 309 unlikely to be directly related to pathophysiology. Some lead terms that attenuate the associations of multiple diseases are related to healthcare utilization (e.g. hospital 310 discharge procedures and "normal diet" text phrases) that are consistent with 311 312 multimorbidity increasing the risk of adverse Covid-19 infection outcomes. Adjustment 313 for electrocardiogram procedures and red cell distribution width laboratory results also 314 attenuated the median point estimate of pulmonary and sleep disease associations with the death outcome more than adjustment for any pulmonary disease. The most notable 315 316 variable to broadly affect associations with death was "blood urea nitrogen" in analyses of women (Figures S12 and S13), which attenuated the point estimates of 12 diseases by 317 >1 and the statistical significance of 13 diseases (p > 0.05), although a smaller effect was 318 319 observed when adjusting for actual median urea nitrogen blood levels collected from -5 to -1 years prior to infection with Covid-19. Adjustment for chronic renal failure and 320 creatinine and estimated glomerular filtration rate laboratory results also support a role 321 for kidney malfunction contributing to pulmonary and sleep disease associations with 322 323 deaths in women.

324 Seventeen of 23 prospectively collected hematology and kidney function
325 laboratory tests were associated with Covid-19 outcome severity at Bonferroni
326 significance (Table S8). Estimated glomerular filtration rate and three red blood cell

Page 17 of 40

Cade et al.: Common pulmonary disease and Covid-19 severity

327 tests were the most significantly associated tests across all models. The laboratory value 328 ranges may reflect physiological changes due to both known and unknown risk factors 329 and may be useful as a general risk stratification measure. These associations were based on a subset of patients with available data, and the associations could possibly 330 also reflect differences in test ordering (e.g. kidney test results may be disproportionally 331 332 missing for patients whose clinicians do not suspect kidney malfunction). Additional 333 follow-up in a broader range of patients is needed to understand the generalization of these results. The prediction models could be readily improved beyond these initial 334 335 exploratory results.

336 The strengths of this study include an extensive sample size of patients diagnosed 337 using PCR testing, which has allowed us to perform sex-, age-, and majority SARS-CoV-338 2 strain date-specific analyses with increased study power. Data were restricted by 339 stringent combined "data floor" and loyalty cohort criteria. We included patients with minimal clinical encounters, notes, diagnoses and evidence of seven clinical facts each 340 collected before the pandemic in order to increase the reliability of our findings in an 341 342 open environment where some patient records may be located in other healthcare systems. All disease and healthcare utilization information was collected prior to Covid-343 19 infection to eliminate potential reverse causation effects. All disease phenotyping 344 included natural language processing to increase the positive predictive value in the face 345 346 of rule-out diagnoses and other potential biases. We extensively studied the potential effect of dozens of pulmonary disease on Covid-19 infection severity, which enabled us 347 to identify the relative importance of pulmonary and disease on infection severity. 348 Weaknesses are also possible with this study design. Despite inclusion of over 349

Cade et al.: Common pulmonary disease and Covid-19 severity

350 37,000 patients, this was a single healthcare system study. While the Mass General 351 Brigham system is comprised of multiple hospitals, our results may or may not 352 completely generalize to other healthcare systems. Loss to follow-up and missing data from out-of-network care may affect our results and those of similarly designed studies. 353 Temporal effects may be confounded by changes in PCR testing availability and use and 354 355 standards of care (e.g. proning). Patients with different financial means may have 356 different levels of healthcare utilization, which may influence the recording of diseases, procedures, and NLP phrases differently. The diseases, procedures, and note terms 357 358 identified by LASSO regression may not completely generalize, despite cross-validation 359 and separate training and testing sets. Individual terms with lower LASSO coefficients, 360 terms that were not individually associated with outcomes (Table S7), or terms that do not affect multiple disease associations may have increased likelihood of site-specific 361 362 associations or may indirectly measure factors not included in our study design. While the majority of the patients had undergone all of the laboratory testing (Table S8), there 363 is likely some selection bias present for patients with clinically relevant diagnoses, 364 365 although some of our results are supported by orthogonal diagnosis and NLP term data 366 (Table S7).

367

368

369 Conclusions

We identified common associations between preexisting pulmonary disease and
Covid-19 infection severity and identified specific diseases with increased relative
importance in our patient sample with multimorbidity. Certain associations are
Page 19 of 40

Cade et al.: Common pulmonary disease and Covid-19 severity

- 373 majority-variant date specific. We also identified several non-pulmonary diseases,
- 374 clinical procedures, and clinical note phrases that potentially mediate or confound these
- 375 associations. These results may aid in further understanding the etiology of a changing
- 376 disease.

Cade et al.: Common pulmonary disease and Covid-19 severity

References

- 1 Shiels MS, Haque AT, Berrington de González A, *et al.* Leading Causes of Death in the US During the COVID-19 Pandemic, March 2020 to October 2021. *JAMA Intern Med* 2022;**182**:883–6. doi:10.1001/jamainternmed.2022.2476
- 2 Armando F, Beythien G, Kaiser FK, *et al.* SARS-CoV-2 Omicron variant causes mild pathology in the upper and lower respiratory tract of hamsters. *Nat Commun* 2022;**13**:3519. doi:10.1038/s41467-022-31200-y
- 3 McMahan K, Giffin V, Tostanoski LH, *et al.* Reduced pathogenicity of the SARS-CoV-2 omicron variant in hamsters. *Med N Y N* 2022;**3**:262-268.e4. doi:10.1016/j.medj.2022.03.004
- 4 Fall A, Eldesouki RE, Sachithanandham J, *et al.* The displacement of the SARS-CoV-2 variant Delta with Omicron: An investigation of hospital admissions and upper respiratory viral loads. *EBioMedicine* 2022;**79**:104008. doi:10.1016/j.ebiom.2022.104008
- 5 Savova GK, Masanz JJ, Ogren PV, *et al.* Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. *J Am Med Inform Assoc JAMIA* 2010;**17**:507–13. doi:10.1136/jamia.2009.001560
- 6 Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res* 2004;**32**:D267-270. doi:10.1093/nar/gkh061
- 7 Liao KP, Sun J, Cai TA, et al. High-throughput multimodal automated phenotyping (MAP) with application to PheWAS. J Am Med Inform Assoc Published Online First: 7 August 2019. doi:10.1093/jamia/ocz066
- 8 Boutin NT, Schecter SB, Perez EF, *et al.* The Evolution of a Large Biobank at Mass General Brigham. *J Pers Med* 2022;**12**:1323. doi:10.3390/jpm12081323
- 9 Karlson EW, Boutin NT, Hoffnagle AG, *et al.* Building the Partners HealthCare Biobank at Partners Personalized Medicine: Informed Consent, Return of Research Results, Recruitment Lessons and Operational Considerations. *J Med* 2016;**6**.
- 10 Weber GM, Adams WG, Bernstam EV, et al. Biases introduced by filtering electronic health records for patients with "complete data." J Am Med Inform Assoc JAMIA 2017;24:1134–41. doi:10.1093/jamia/ocx071
- 11 Robinson JR, Wei W-Q, Roden DM, et al. Defining Phenotypes from Clinical Data to Drive Genomic Research. Annu Rev Biomed Data Sci 2018;1:69–92. doi:10.1146/annurev-biodatasci-080917-013335
- 12 Wei W-Q, Bastarache LA, Carroll RJ, *et al.* Evaluating phecodes, clinical classification software, and ICD-9-CM codes for phenome-wide association studies in the electronic health record. *PloS One* 2017;**12**:e0175508. doi:10.1371/journal.pone.0175508
- 13 Wu P, Gifford A, Meng X, *et al.* Mapping ICD-10 and ICD-10-CM Codes to Phecodes: Workflow Development and Initial Evaluation. *JMIR Med Inform* 2019;7:e14325. doi:10.2196/14325

Cade et al.: Common pulmonary disease and Covid-19 severity

- 14 Cade BE, Dashti HS, Hassan SM, *et al.* Sleep Apnea and COVID-19 Mortality and Hospitalization. *Am J Respir Crit Care Med* 2020;**202**:1462–4. doi:10.1164/rccm.202006-2252LE
- 15 Jones SE, Maisha FI, Strausz SJ, *et al.* The public health impact of poor sleep on severe COVID-19, influenza and upper respiratory infections. *MedRxiv Prepr Serv Health Sci* 2022;:2022.02.16.22271055. doi:10.1101/2022.02.16.22271055
- 16 Davis HE, Assaf GS, McCorkell L, *et al.* Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *EClinicalMedicine* 2021;**38**:101019. doi:10.1016/j.eclinm.2021.101019
- 17 Strasser ZH, Greifer N, Hadavand A, *et al.* Estimates of SARS-CoV-2 Omicron BA.2 Subvariant Severity in New England. *JAMA Netw Open* 2022;5:e2238354. doi:10.1001/jamanetworkopen.2022.38354

Cade *et al.*: Common pulmonary disease and Covid-19 severity Figure Legends

379 Figure 1. Effect of lead LASSO terms and laboratory results on disease

380 associations with Covid-19 death outcome.

- 381 The y axis includes all pulmonary and sleep diseases associated with death in the
- 382 combined sex analyses (Table 2). Lead LASSO terms (Table S6) and laboratory results
- 383 (Table S8) were included individually as covariates in each of these disease associations.
- 384 The odds ratio difference when including versus not including an adjusted covariate are
- 385 shown as circles, with darker red indicating an attenuated odds ratio for the Y-axis
- 386 disease association with death when adjusting for the X-axis covariate. The X-axis
- 387 covariates are ordered by median value of point estimates considering all of the Y-axis
- 388 diseases. Additional visualizations considering the three outcomes presented as point
- 389 estimate and p-value differences are provided in Figures S1 S15. Full results with
- 390 sortable values are provided in Table S9.

Cade et al.: Common pulmonary disease and Covid-19 severity

Tables

Table 1. Sample characteristics.

	All	Death Outco me	Comp osite Outco me	Inpatie nt Outco me	Breakt hrough Infectio n	Pre- Delta variant majorit y dates	Delta variant majorit y dates	Omicro n variant majorit y dates
N	37,02 0	924	1,166	4,338	15,472	18,154	4,916	13,950
Women	22,81 8 (61.6)	443 (47.9)	558 (47.9)	2,359 (54.4)	9,614 (62.1)	11,026 (60.7)	2,925 (59.5)	8,867 (63.6)
Age	54.1 (25.3)	78.5 (17.7)	68.8 (20.8)	69.2 (23.4)	56.0 (25.0)	53.7 (24.9)	54.3 (25.5)	54.8 (25.8)
BMI	28.3 (8.0)	26.6 (8.5)	29.2 (8.4)	28.8 (8.6)	27.7 (8.1)	28.7 (7.9)	28.1 (7.9)	27.7 (8.2)
Race/ ethnicity								
European American	26,65 9 (72.0)	768 (83.1)	782 (67.1)	2,996 (69.1)	12,422 (80.3)	11,951 (65.8)	4,021 (81.8)	10,687 (76.6)
Hispanic/ Latino American	3,764 (10.2)	55 (6.0)	155 (13.3)	519 (12.0)	784 (5.1)	2,716 (15.0)	201 (4.1)	847 (6.1)
African American	3,176 (8.6)	55 (6.0)	129 (11.1)	495 (11.4)	978 (6.3)	1,731 (9.5)	343 (7.0)	1,102 (7.9)
Other	3,421 (9.2)	46 (5.0)	100 (8.6)	328 (7.6)	1,288 (8.3)	1,756 (9.7)	351 (7.1)	1,314 (9.4)
Pulmonar y and sleep diseases	2 (3)	4 (6)	3 (5)	3 (5)	2 (3)	2 (3)	2 (3)	2 (3)
Acute bronchitis and bronchioli tis	2,479 (6.7)	92 (10.0)	129 (11.1)	446 (10.3)	992 (6.4)	1,259 (6.9)	323 (6.6)	897 (6.4)

	<u> </u>	<u>ae et at.: (</u>	ommon pu	unnonary di	sease and U	ovia-19 seve	rity	
Acute pharyngiti s	5,276 (14.3)	41 (4.4)	101 (8.7)	368 (8.5)	2,240 (14.5)	2,429 (13.4)	715 (14.5)	2,132 (15.3)
Acute sinusitis	2,434 (6.6)	31 (3.4)	59 (5.1)	216 (5.0)	1,152 (7.5)	1,088 (6.0)	320 (6.5)	1,026 (7.4)
Acute upper respirator y infections of multiple or unspecifie d sites	5,204 (14.1)	89 (9.6)	146 (12.5)	494 (11.4)	2,345 (15.2)	2,361 (13.0)	716 (14.6)	2,127 (15.3)
Allergy	8,596 (23.2)	193 (20.9)	283 (24.3)	967 (22.3)	3,841 (24.8)	3,924 (21.6)	1,162 (23.6)	3,510 (25.2)
ARDS	300 (0.8)	32 (3.5)	38 (3.3)	123 (2.8)	130 (0.8)	150 (0.8)	39 (0.8)	111 (0.8)
Asphyxia and hypoxemi a	531 (1.4)	79 (8.6)	68 (5.8)	276 (6.4)	215 (1.4)	270 (1.5)	61 (1.2)	200 (1.4)
Asthma	6,103 (16.5)	134 (14.5)	225 (19.3)	777 (17.9)	2,590 (16.7)	2,838 (15.6)	821 (16.7)	2,444 (17.5)
Bacterial infection NOS	1,324 (3.6)	87 (9.4)	79 (6.8)	287 (6.6)	500 (3.2)	676 (3.7)	177 (3.6)	471 (3.4)
Bronchitis	2,058 (5.6)	102 (11.0)	109 (9.4)	395 (9.1)	943 (6.1)	937 (5.2)	290 (5.9)	831 (6.0)
Chronic airway obstructio n	1,165 (3.2)	167 (18.1)	144 (12.4)	504 (11.6)	505 (3.3)	576 (3.2)	143 (2.9)	446 (3.2)
Chronic bronchitis	549 (1.5)	76 (8.2)	67 (5.8)	261 (6.0)	226 (1.5)	264 (1.5)	57 (1.2)	228 (1.6)
Chronic pharyngiti s and nasophary	1,780 (4.8)	47 (5.1)	45 (3.9)	207 (4.8)	828 (5.4)	804 (4.4)	249 (5.1)	727 (5.2)

Cade et al.: Common	pulmonary	disease and	Covid-19	severity
---------------------	-----------	-------------	----------	----------

	<u> </u>	ue et util (ommon p	annonai y u	sease and e	0114 19 5010	<u> </u>	
ngitis								
Chronic sinusitis	2,051 (5.5)	39 (4.2)	56 (4.8)	221 (5.1)	981 (6.3)	903 (5.0)	282 (5.7)	866 (6.2)
Chronic tonsillitis and adenoiditi s	608 (1.6)	5 (0.5)	11 (0.9)	47 (1.1)	263 (1.7)	262 (1.4)	99 (2.0)	247 (1.8)
Cough	7,152 (19.3)	272 (29.4)	324 (27.8)	1,202 (27.7)	3,247 (21.0)	3,326 (18.3)	945 (19.2)	2,881 (20.7)
Emphyse ma	716 (1.9)	98 (10.6)	97 (8.3)	311 (7.2)	326 (2.1)	332 (1.8)	95 (1.9)	289 (2.1)
Empyema and pneumoth orax	602 (1.6)	62 (6.7)	53 (4.6)	197 (4.5)	284 (1.8)	269 (1.5)	92 (1.9)	241 (1.7)
Epistaxis or throat hemorrha ge	1,185 (3.2)	80 (8.7)	64 (5.5)	226 (5.2)	514 (3.3)	564 (3.1)	168 (3.4)	453 (3.3)
Hypersom nia	584 (1.6)	17 (1.8)	24 (2.1)	86 (2.0)	265 (1.7)	261 (1.4)	73 (1.5)	250 (1.8)
Influenza	2,442 (6.6)	100 (10.8)	125 (10.7)	455 (10.5)	1,035 (6.7)	1,195 (6.6)	284 (5.8)	963 (6.9)
Insomnia	3,398 (9.2)	111 (12.0)	140 (12.0)	488 (11.3)	1,605 (10.4)	1,503 (8.3)	420 (8.5)	1,475 (10.6)
Lung cancer	329 (0.9)	51 (5.5)	26 (2.2)	130 (3.0)	163 (1.1)	147 (0.8)	38 (0.8)	144 (1.0)
Lymphang ioleiomyo matosis and others	182 (0.5)	33 (3.6)	22 (1.9)	82 (1.9)	97 (0.6)	74 (0.4)	24 (0.5)	84 (0.6)
Other diseases of lung	1,279 (3.5)	138 (14.9)	112 (9.6)	446 (10.3)	546 (3.5)	646 (3.6)	147 (3.0)	486 (3.5)
Other dyspnea	2,931 (7.9)	228 (24.7)	221 (19.0)	828 (19.1)	1,366 (8.8)	1,336 (7.4)	382 (7.8)	1,213 (8.7)
Other upper	3,664 (9.9)	92 (10.0)	114 (9.8)	471 (10.9)	1,819 (11.8)	1,508 (8.3)	488 (9.9)	1,668 (12.0)

Cade *et al.*: Common pulmonary disease and Covid-19 severity

	Cu		Johnnon pe	annonar y ui	scase and c	0111 19 SCIC	i icy	
respirator y disease								
Painful respiratio n	2,176 (5.9)	77 (8.3)	102 (8.8)	412 (9.5)	903 (5.8)	1,044 (5.8)	275 (5.6)	857 (6.1)
Parasomni a	199 (0.5)	7 (0.8)	6 (0.5)	33 (0.8)	82 (0.5)	99 (0.6)	29 (0.6)	71 (0.5)
Pleurisy; pleural effusion	1,017 (2.8)	145 (15.7)	120 (10.3)	421 (9.7)	457 (3.0)	483 (2.7)	126 (2.6)	408 (2.9)
Pneumoni a	2,375 (6.4)	224 (24.2)	212 (18.2)	779 (18.0)	1,065 (6.9)	1,122 (6.2)	299 (6.1)	954 (6.8)
Postnasal drip	1,318 (3.6)	43 (4.7)	47 (4.0)	188 (4.3)	665 (4.3)	557 (3.1)	177 (3.6)	584 (4.2)
Pulmonar y collapse; interstitial and compensa tory emphyse ma	1,198 (3.2)	153 (16.6)	124 (10.6)	505 (11.6)	553 (3.6)	579 (3.2)	146 (3.0)	473 (3.4)
Pulmonar y congestion and hypostasis	722 (2.0)	123 (13.3)	111 (9.5)	398 (9.2)	304 (2.0)	368 (2.0)	86 (1.8)	268 (1.9)
Pulmonar y embolism and infarction, acute	649 (1.8)	55 (6.0)	58 (5.0)	200 (4.6)	287 (1.9)	300 (1.7)	93 (1.9)	256 (1.8)
Pulmonar y fibrosis	266 (0.7)	42 (4.6)	30 (2.6)	110 (2.5)	120 (0.8)	132 (0.7)	23 (0.5)	111 (0.8)
Pulmonar y hypertensi on	366 (1.0)	60 (6.5)	54 (4.6)	175 (4.0)	166 (1.1)	175 (1.0)	49 (1.0)	142 (1.0)
Respirator	471	83	86 (7.4)	252	189 (1.2)	238 (1.3)	55 (1.1)	178 (1.3)

Cade et al.: Common pulmonary disease and Covid-19 severity

y failure	(1.3)	(9.0)		(5.8)			,11 (y	
Restless legs syndrome	456 (1.2)	26 (2.8)	29 (2.5)	91 (2.1)	215 (1.4)	210 (1.2)	56 (1.1)	190 (1.4)
Septal deviations / turbinate hypertrop hy	1,321 (3.6)	22 (2.4)	23 (2.0)	94 (2.2)	643 (4.2)	543 (3.0)	204 (4.2)	574 (4.1)
Shortness of breath	3,729 (10.1)	273 (29.6)	270 (23.2)	1,063 (24.5)	1,729 (11.2)	1,663 (9.2)	463 (9.4)	1,603 (11.5)
Sleep apnea	4,240 (11.5)	166 (18.0)	215 (18.4)	717 (16.5)	1,946 (12.6)	1,962 (10.8)	596 (12.1)	1,682 (12.1)
Sleep disorders	1,260 (3.4)	41 (4.4)	43 (3.7)	214 (4.9)	541 (3.5)	616 (3.4)	160 (3.3)	484 (3.5)
Solitary pulmonar y nodule	1,542 (4.2)	112 (12.1)	101 (8.7)	379 (8.7)	756 (4.9)	667 (3.7)	220 (4.5)	655 (4.7)
Symptoms involving respirator y system and other chest symptoms	873 (2.4)	62 (6.7)	62 (5.3)	216 (5.0)	390 (2.5)	419 (2.3)	108 (2.2)	346 (2.5)
Throat pain	822 (2.2)	13 (1.4)	20 (1.7)	78 (1.8)	347 (2.2)	375 (2.1)	100 (2.0)	347 (2.5)
Tuberous sclerosis complex and others	190 (0.5)	4 (0.4)	3 (0.3)	32 (0.7)	81 (0.5)	99 (0.6)	22 (0.5)	69 (0.5)
Viral infection	2,596 (7.0)	42 (4.6)	62 (5.3)	279 (6.4)	976 (6.3)	1,323 (7.3)	301 (6.1)	972 (7.0)
Viremia, NOS	713 (1.9)	17 (1.8)	29 (2.5)	111 (2.6)	295 (1.9)	321 (1.8)	84 (1.7)	308 (2.2)
Voice disturbanc	867 (2.3)	29 (3.1)	24 (2.1)	133 (3.1)	441 (2.9)	370 (2.0)	100 (2.0)	397 (2.9)

Cade et al.: Common pulmonary disease and Covid-19 severity

	Ca		Johnnon pu	innonai y u	scase and c	0viu=19 scvc	-1 ILY	
Wheezing	1,556 (4.2)	85 (9.2)	90 (7.7)	383 (8.8)	669 (4.3)	726 (4.0)	202 (4.1)	628 (4.5)

Cade et al.: Common pulmonary disease and Covid-19 severity

The composite outcome consisted of mechanical ventilation or ICU admission. A breakthrough infection was defined as having occurred >30 days after a second vaccine administration. Values in the three outcome and the breakthrough columns are calculated on the subset of patients who experienced that outcome. Transition dates from the original to the Delta and Omicron variants were set at 7/3/2021 and 12/17/2021 respectively. ICD and PheCode definitions of each disease are provided in Table S1. All values are presented as count (percentage); apart from Age, BMI, and pulmonary and sleep diseases; which are presented as median (interquartile range).

Cade *et al.*: Common pulmonary disease and Covid-19 severity Table 2. Associations between preexisting disease and Covid-19 infection severity.

		Mechanical	
		ventilation or ICU	Inpatient
		admission	admission
Existing Disorder	Death Outcome	Outcome	Outcome
Acute bronchitis and			
bronchiolitis	1.25 (0.99 – 1.59)	1.43 (1.18 – 1.74)†	1.42 (1.27 – 1.60)†
Acute pharyngitis	0.78 (0.56 - 1.08)	1.05 (0.85 - 1.31)	0.96 (0.85 - 1.08)
Acute sinusitis	0.69 (0.48 – 1.01)	0.94 (0.72 – 1.24)	0.83 (0.72 - 0.97)*
Acute upper			
respiratory infections			
of multiple or			
unspecified sites	0.88 (0.70 - 1.12)	1.12 (0.93 – 1.35)	0.92 (0.83 – 1.02)
Allergy	1.13 (0.95 - 1.34)	1.29 (1.12 - 1.49)*	1.10 (1.01 – 1.19)*
		2.95 (2.03 –	3.90 (3.01 –
ARDS	2.48 (1.62 - 3.81)†	4.29)†	5.07)†
Asphyxia and			5.35 (4.41 -
hypoxemia	2.93 (2.21 - 3.89)†	2.36 (1.78 - 3.13)†	6.49)†
Asthma	1.18 (0.97 - 1.44)	1.40 (1.20 - 1.64)†	1.28 (1.17 - 1.40)†
Bacterial infection	2.68 (2.06 -		2.30 (1.98 -
NOS	3.48) †	1.96 (1.53 - 2.52)†	2.68)†
	• • ·		1.40 (1.24 –
Bronchitis	1.45 (1.16 – 1.83)*	1.38 (1.12 – 1.70)*	1.59)†
Chronic airway			3.18 (2.79 -
obstruction	2.70 (2.22 - 3.28)†	2.42 (1.99 – 2.96)†	3.63)†
			3.53 (2.93 -
Chronic bronchitis	2.30 (1.75 - 3.03)†	2.11 (1.59 - 2.78)†	4.25)†
Chronic pharyngitis			
and nasopharyngitis	0.88 (0.64 - 1.21)	0.72 (0.53 - 0.99)*	0.89 (0.76 – 1.04)
Chronic sinusitis	0.84 (0.60 - 1.18)	0.97 (0.73 – 1.28)	0.97 (0.84 - 1.13)
Chronic tonsillitis and			
adenoiditis	1.06 (0.43 – 2.63)	1.05 (0.57 – 1.94)	1.23 (0.89 – 1.69)
Cough	1.25 (1.07 - 1.46)*	1.31 (1.14 - 1.51)†	1.39 (1.29 - 1.51)†
		2.90 (2.29 -	3.20 (2.71 –
Emphysema	2.39 (1.87 - 3.05)†	3.68) †	3.77)†
Empyema and		-	2.65 (2.19 -
pneumothorax	2.38 (1.76 - 3.21)†	2.04 (1.50 - 2.76)†	3.21)†
Epistaxis or throat			
hemorrhage	1.88 (1.45 - 2.45)†	1.33 (1.01 – 1.74)*	1.35 (1.15 - 1.59)†
Hypersomnia	1.47 (0.88 - 2.46)	1.27 (0.83 – 1.94)	1.27 (1.00 - 1.63)
			1.68 (1.49 –
Influenza	1.47 (1.16 – 1.85)*	1.54 (1.26 - 1.88)†	1.89)†
Insomnia	1.09 (0.88 - 1.35)	1.26 (1.04 - 1.52)*	1.09 (0.98 - 1.22)

Cade e	<i>t al.</i> : Common pulmonary	y disease and Covid-19 sev	verity
		Mechanical	
		ventilation or ICU	Inpatient
		admission	admission
Existing Disorder	Death Outcome	Outcome	Outcome
			2.8 7 (2.25 –
Lung cancer	3.23 (2.31 - 4.52)†	1.55 (1.02 – 2.37)*	3.66) †
Lymphangioleiomyo			3.82 (2. 77 –
matosis	4.57 (2.96 – 7.04) †	2.54 (1.5 7 – 4.11)†	5.26) †
			2.33 (2.04 –
Other diseases of lung	2.13 (1.73 - 2.63)†	1.66 (1.34 - 2.05)†	2.65) †
			2.33 (2.12 –
Other dyspnea	2.28 (1.92 - 2.70)†	1.85 (1.58 - 2.18)†	2.57)†
Other upper			
respiratory disease	1.13 (0.89 – 1.42)	1.09 (0.89 – 1.34)	1.23 (1.10 – 1.37)†
Painful respiration	1.23 (0.96 – 1.60)	1.33 (1.07 – 1.65)*	1.58 (1.40 – 1.79)†
Parasomnia	1.36 (0.61 – 3.03)	0.85 (0.37 – 1.96)	1.39 (0.93 - 2.09)
Pleurisy; pleural			3.43 (2.97 -
effusion	2.89 (2.33 - 3.58)†	2.45 (1.97 - 3.04)†	3.96)†
			2.85 (2.57 -
Pneumonia	2.16 (1.81 - 2.58)†	2.13 (1.80 - 2.51)†	3.16)†
Postnasal drip	0.89 (0.64 - 1.24)	1.00 (0.73 - 1.35)	1.02(0.86 - 1.20)
Pulmonary collapse;			
interstitial and			
compensatory			3.65 (3.19 –
emphysema	2.59 (2.10 - 3.18)†	2.05 (1.66 - 2.54)†	4.17)†
Pulmonary			
congestion and			5.85 (4.95 -
hypostasis	3.55 (2.81 - 4.48)†	3.04 (2.41 - 3.82)†	6.91)†
Pulmonary embolism			2.26 (1.87 –
and infarction, acute	1.87 (1.36 - 2.56)†	1.91 (1.42 – 2.56)†	2.72)†
			3.03 (2.31 –
Pulmonary fibrosis	3.26 (2.22 - 4.78)†	2.09 (1.39 – 3.16)*	3.97)†
Pulmonary		2.79 (2.03 –	3.67 (2.92 –
hypertension	3.15 (2.28 – 4.33) †	3.83) †	4.61) †
			6.52 (5.31 –
Respiratory failure	4.69 (3.54 - 6.21)†	4.44 (3.41 - 5.78)†	8.01) †
Restless legs			
syndrome	1.58 (1.02 – 2.44)*	1.61 (1.08 – 2.41)*	1.38 (1.07 – 1.77)*
Septal deviations /			0.64 (0.52 -
turbinate hypertrophy	0.85 (0.54 - 1.32)	0.63 (0.41 – 0.96)*	0.80) †
			2.48 (2.27 –
Shortness of breath	2.04 (1.74 – 2.40) †	1.81 (1.56 – 2.10) †	2.71)†
Sleep apnea	$1.68(1.38 - 2.03)^{\dagger}$	1.32 (1.12 – 1.56)*	1.27 (1.15 – 1.40) †
Sleep disorders	1.07 (0.76 – 1.50)	0.88 (0.64 - 1.22)	1.35 (1.15 – 1.59) †
Solitary pulmonary	$1.64(1.31 - 2.04)^{\dagger}$	1.49 (1.20 - 1.86)*	$1.56(1.37 - 1.77)^{\dagger}$

Cade e	et al.: Common pulmonary	v disease and Covid-19 sev	erity
		Mechanical	
		ventilation or ICU	Inpatient
		admission	admission
Existing Disorder	Death Outcome	Outcome	Outcome
nodule			
Symptoms involving			
respiratory system			
and other chest			1.98 (1.66 –
symptoms	2.27 (1.69 - 3.04)†	1.90 (1.44 – 2.51)†	2.35) †
Throat pain	0.94 (0.53 – 1.67)	0.92 (0.58 – 1.46)	0.90 (0.70 – 1.16)
Tuberous Sclerosis			2.51 (1.66 –
Complex	1.86 (0.66 – 5.28)	0.71 (0.22 – 2.27)	3.80) †
Viral infection	1.47 (1.05 - 2.05)*	1.16 (0.89 – 1.52)	1.51 (1.31 – 1.74)†
Viremia, NOS	1.30 (0.77 – 2.19)	1.52 (1.03 – 2.25)*	1.71 (1.37 – 2.14) †
Voice disturbance	0.88 (0.59 - 1.31)	0.71 (0.47 – 1.08)	1.03 (0.85 – 1.26)
			2.09 (1.83 -
Wheezing	1.76 (1.37 – 2.26) †	1.44 (1.14 – 1.81)*	2.38) †

Odds ratios with 95% confidence intervals are shown. All association calculations adjusted for age, sex, race/ethnicity, obesity category, and breakthrough infection status. *: p < 0.05, †: $p < 3.3 \times 10^{-4}$ (Bonferroni significance).

Cade *et al.*: Common pulmonary disease and Covid-19 severity Table 3. Sex-stratified associations between preexisting disease and death.

Existing Disorder	Men	Women	P Sex Interaction
Acute bronchitis and bronchiolitis	1.10 (0.77 – 1.59)	1.35 (0.99 – 1.84)	0.366
Acute pharyngitis	0.52 (0.28 – 0.97)*	0.96 (0.65 – 1.42)	0.104
Acute sinusitis	0.33 (0.14 – 0.77)*	0.94 (0.61 – 1.43)	0.035
Acute upper respiratory infections of multiple or unspecified sites	0.67 (0.45 – 1.00)*	1.05 (0.78 – 1.40)	0.063
Allergy	1.01 (0.77 – 1.32)	1.23 (0.98 – 1.54)	0.274
ARDS	2.76 (1.54 – 4.95)*	2.15 (1.14 – 4.07)*	0.621
Asphyxia and hypoxemia	2.68 (1. 77 – 4.05) †	3.11 (2.11 – 4.57)†	0.486
Asthma	0.85 (0.61 – 1.19)	1.42 (1.11 – 1.83)*	0.012
Bacterial infection NOS	2.35 (1.57 – 3.53)†	2.93 (2.07 – 4.13)†	0.402
Bronchitis	1.22 (0.86 – 1.74)	1.68 (1.25 – 2.27)*	0.167
Chronic airway obstruction	2.56 (1.93 – 3.40)†	2.84 (2.16 – 3.72)†	0.667
Chronic bronchitis	1.89 (1.23 – 2.89)*	2.64 (1.84 – 3.81)†	0.241
Chronic pharyngitis and nasopharyngitis	0.91 (0.59 – 1.41)	0.86 (0.54 – 1.38)	0.852
Chronic sinusitis	1.14 (0.73 – 1.77)	0.60 (0.35 – 1.02)	0.071
Chronic tonsillitis and adenoiditis	1.35 (0.41 – 4.42)	0.82 (0.20 – 3.41)	0.598
Cough	1.12 (0.89 – 1.40)	1.37 (1.11 – 1.70)*	0.109
Emphysema	2.32 (1.64 – 3.28) †	2.41 (1.70 – 3.42)†	0.976
Empyema and pneumothorax	2.68 (1.85 – 3.88) †	1.87 (1.10 – 3.19)*	0.221
Epistaxis or throat hemorrhage	1.87 (1.30 – 2.69)*	1.89 (1.29 – 2.76)*	0.795
Hypersomnia	1.68 (0.90 – 3.13)	1.21 (0.48 – 3.03)	0.607
Influenza	1.41 (0.99 – 2.00)	1.50 (1.10 – 2.06)*	0.629

Cade et al.: C	ommon pulmonary disease	and Covid-19 severity	1
Existing Disorder	Men	Women	P Sex Interaction
Insomnia	1.25 (0.90 – 1.72)	0.99 (0.74 – 1.33)	0.353
Lung cancer	2.39 (1.44 - 3.97)*	4.14 (2.63 – 6.51)†	0.110
Lymphangioleiomyomatosi s	5.15 (2.93 – 9.04) †	4.07 (2.05 – 8.09)†	0.828
Other diseases of lung	1.86 (1.38 – 2.52)†	2.39 (1.79 – 3.18)†	0.210
Other dyspnea	2.10 (1.65 – 2.67) †	2.44 (1.91 – 3.10)†	0.203
Other upper respiratory disease	1.05 (0.75 – 1.46)	1.22 (0.88 – 1.69)	0.470
Painful respiration	1.37 (0.94 – 1.99)	1.09 (0.76 – 1.55)	0.381
Parasomnia	2.22 (0.88 - 5.58)	0.42 (0.06 - 3.08)	0.143
Pleurisy; pleural effusion	2.60 (1.94 – 3.47) †	3.27 (2.38 – 4.49)†	0.253
Pneumonia	1.98 (1.55 – 2.55)†	2.33 (1.82 – 3.00)†	0.387
Postnasal drip	1.08 (0.67 – 1.72)	0.77 (0.48 – 1.24)	0.399
Pulmonary collapse; interstitial and compensatory emphysema	2.61 (1.99 – 3.42) †	2.51 (1.81 – 3.46)†	0.935
Pulmonary congestion and hypostasis	4.31 (3.11 – 5.96) †	2.87 (2.05 – 4.02)†	0.170
Pulmonary embolism and infarction, acute	1.93 (1.24 – 3.00)*	1.74 (1.11 – 2.73)*	0.899
Pulmonary fibrosis	3.29 (1.90 – 5.71) †	3.35 (1.97 – 5.70)†	0.870
Pulmonary hypertension	3.62 (2.21 - 5.93)†	2.79 (1.84 – 4.25)†	0.570
Respiratory failure	4.12 (2. 77 – 6.13) †	5.35 (3.58 – 7.99)†	0.364
Restless legs syndrome	2.08 (1.11 – 3.90)*	1.28 (0.70 – 2.35)	0.377
Septal deviations / turbinate hypertrophy	1.06 (0.63 – 1.79)	0.53 (0.22 – 1.31)	0.191

Cade et al.: Commor	pulmonary	disease	and Covid-	19 severity
---------------------	-----------	---------	------------	-------------

Existing Disorder	Men	Women	P Sex
			Interaction
Shortness of breath	1.84 (1.46 – 2.32)†	2.22 (1.78 – 2.78)†	0.123
Sleep apnea	1.68 (1.31 – 2.15)†	1.79 (1.32 – 2.44)*	0.306
Sleep disorders	1.08 (0.66 – 1.75)	1.07 (0.66 – 1.74)	0.914
Solitary pulmonary nodule	1.56 (1.14 – 2.14)*	1.65 (1.21 – 2.25)*	0.754
Symptoms involving respiratory system and other chest symptoms	2.13 (1.38 – 3.30)*	2.37 (1.58 – 3.53)†	0.711
Throat pain	0.56 (0.17 – 1.80)	1.20 (0.61 – 2.33)	0.244
Tuberous Sclerosis Complex	0.00 (0.00 – Inf)	3.13 (1.06 – 9.28)*	0.945
Viral infection	1.36 (0.82 – 2.25)	1.54 (0.99 – 2.42)	0.746
Viremia, NOS	0.71 (0.28 – 1.84)	1.85 (0.99 – 3.46)	0.093
Voice disturbance	0.96 (0.54 – 1.70)	0.82 (0.48 - 1.43)	0.747
Wheezing	2.03 (1.41 – 2.93) †	1.53 (1.08 – 2.17)*	0.441

Cade et al.: Common pulmonary disease and Covid-19 severity

Associations with a death outcome are shown using odds ratios with 95% confidence intervals. Results from all three outcomes can be found in Table S2. All association calculations adjusted for age, sex, race/ethnicity, obesity category, and breakthrough infection status. *: p < 0.05, †: $p < 1.6 \times 10^{-4}$ (Bonferroni significance).

Cade et al.: Common pulmonary disease and Covid-19 severity
Table 4. Date-specific associations between preexisting disease and death

Existing disease	Pre-Delta variant majority dates	Delta variant majority dates	Omicron variant majority dates
Acute bronchitis and bronchiolitis	1.25 (0.94 – 1.67)	1.08 (0.52 – 2.25)	1.25 (0.75 – 2.08)
Acute pharyngitis	0.79 (0.52 – 1.19)	0.53 (0.16 – 1.73)	0.91 (0.50 – 1.67)
Acute sinusitis	0.65 (0.40 - 1.06)	0.65 (0.23 - 1.82)	0.81 (0.39 – 1.68)
Acute upper respiratory infections of multiple or unspecified sites	0.89 (0.66 – 1.20)	1.31 (0.73 – 2.35)	0.71 (0.43 – 1.16)
Allergy	1.12 (0.90 – 1.40)	1.02 (0.60 – 1.71)	1.20 (0.85 – 1.70)
ARDS	2.55 (1.47 - 4.40)*	2.77 (0.88 – 8.71)	2.33 (0.96 - 5.63)
Asphyxia and hypoxemia	3.39 (2.40 – 4.80)†	1.53 (0.50 – 4.65)	2.52 (1.40 - 4.53)*
Asthma	1.13 (0.88 – 1.46)	1.24 (0.69 – 2.22)	1.34 (0.91 – 1.97)
Bacterial infection NOS	3.00 (2.16 – 4.16)†	3.91 (1.89 – 8.05)*	1.63 (0.90 – 2.95)
Bronchitis	1.45 (1.09 – 1.94)*	1.43 (0.73 – 2.80)	1.43 (0.90 – 2.26)
Chronic airway obstruction	2.59 (2.02 – 3.33)†	2.19 (1.15 – 4.18)*	3.47 (2.39 – 5.04)†
Chronic bronchitis	2.26 (1.58 – 3.23)†	1.84 (0.69 – 4.91)	3.07 (1.87 – 5.03)†
Chronic pharyngitis and nasopharyngitis	0.86 (0.57 – 1.30)	1.25 (0.58 – 2.71)	0.79 (0.40 – 1.56)
Chronic sinusitis	0.90 (0.59 – 1.39)	0.99 (0.42 – 2.35)	0.64 (0.30 – 1.35)
Chronic tonsillitis and adenoiditis	1.90 (0.74 – 4.84)	N/A	N/A
Cough	1.15 (0.94 – 1.40)	1.71 (1.09 – 2.69)*	1.30 (0.95 - 1.78)
Emphysema	2.35 (1.71 – 3.24)†	1.91 (0.87 – 4.22)	2.86 (1.83 – 4.48)†
Empyema and pneumothorax	2.83 (1.95 – 4.12)†	1.24 (0.45 – 3.42)	1.94 (1.03 – 3.65)*
Epistaxis or throat hemorrhage	1.97 (1.42 – 2.74)†	1.63 (0.71 – 3.72)	1.69 (0.99 – 2.87)

Cade e	<i>t al.</i> : Common pulmonal	ry disease and Covid-19 se	verity
Existing disease	Pre-Delta variant majority dates	Delta variant majority dates	Omicron variant majority dates
Hypersomnia	1.23 (0.61 – 2.50)	1.23 (0.25 – 6.10)	2.30 (0.98 - 5.40)
Influenza	1.72 (1.29 – 2.30)*	1.57 (0.79 - 3.13)	0.90 (0.53 – 1.55)
Insomnia	1.17 (0.89 – 1.54)	1.61 (0.89 – 2.92)	0.78 (0.49 – 1.22)
Lung cancer	3.44 (2.20 – 5.38)†	1.52 (0.43 – 5.45)	3.75 (2.11 - 6.65)†
Lymphangioleiomyo matosis	4.76 (2.59 – 8.72)†	2.08 (0.45 - 9.72)	5.42 (2.69 – 10.93)†
Other diseases of lung	2.12 (1.63 – 2.75)†	1.27 (0.61 – 2.65)	2.69 (1.81 – 3.98)†
Other dyspnea	2.20 (1.76 – 2.74)†	2.79 (1.71 – 4.54) †	2.26 (1.61 - 3.17)†
Other upper respiratory disease	1.23 (0.92 – 1.66)	1.15 (0.59 – 2.24)	0.98 (0.62 – 1.56)
Painful respiration	1.19 (0.86 – 1.65)	0.69 (0.27 – 1.77)	1.64 (1.02 – 2.66)*
Parasomnia	1.90 (0.77 – 4.71)	N/A	0.81 (0.11 – 6.14)
Pleurisy; pleural effusion	2.89 (2.18 – 3.81)†	4.44 (2.37 – 8.31)†	2.61 (1.73 - 3.96)†
Pneumonia	2.38 (1.91 – 2.97)†	2.50 (1.47 – 4.26)*	1.59 (1.10 – 2.30)*
Postnasal drip	0.85 (0.54 - 1.32)	1.45 (0.64 – 3.30)	0.81 (0.42 – 1.54)
Pulmonary collapse; interstitial and compensatory emphysema	2.33 (1.78 – 3.05)†	3.13 (1.70 – 5.78)*	3.00 (2.03 – 4.42)†
Pulmonary congestion and hypostasis	3.43 (2.54 – 4.63)†	5.37 (2.80 – 10.29)†	3.06 (1.90 – 4.94)†
Pulmonary embolism and infarction, acute	1.68 (1.12 – 2.53)*	1.17 (0.44 – 3.12)	2.92 (1.66 – 5.16)*
Pulmonary fibrosis	3.63 (2.25 – 5.86)†	1.33 (0.16 – 10.78)	2.96 (1.41 – 6.20)*
Pulmonary hypertension	3.25 (2.17 – 4.88)†	3.82 (1.63 – 8.98)*	2.33 (1.16 – 4.69)*
Respiratory failure	5.62 (3.92 – 8.07)†	5.36 (2.34 – 12.27)†	3.51 (1.99 – 6.17)†

Cade *et al.*: Common pulmonary disease and Covid-19 severity

Existing discasse Dro Dolto variant Dolto variant Originary				
Existing disease	majority dates	majority dates	majority dates	
Restless legs syndrome	1.65 (0.96 – 2.84)	0.71 (0.09 – 5.41)	1.83 (0.83 – 4.07)	
Septal deviations / turbinate hypertrophy	1.22 (0.72 – 2.05)	0.57 (0.14 – 2.40)	0.42 (0.13 – 1.33)	
Shortness of breath	2.09 (1.71 – 2.57)†	1.75 (1.06 – 2.87)*	2.16 (1.58 – 2.96)†	
Sleep apnea	1.83 (1.44 – 2.33)†	1.64 (0.94 – 2.88)	1.35 (0.89 – 2.04)	
Sleep disorders	1.39 (0.94 – 2.06)	0.48 (0.11 – 2.03)	0.57 (0.23 – 1.41)	
Solitary pulmonary nodule	1.47 (1.09 – 1.97)*	1.37 (0.71 – 2.63)	2.24 (1.52 – 3.29)†	
Symptoms involving respiratory system and other chest symptoms	2.31 (1.57 – 3.39)†	2.54 (1.16 – 5.57)*	2.07 (1.14 – 3.76)*	
Throat pain	0.78 (0.33 - 1.82)	1.44 (0.33 – 6.31)	1.19 (0.47 – 3.03)	
Tuberous Sclerosis Complex	1.85 (0.55 – 6.25)	N/A	1.92 (0.23 – 15.79)	
Viral infection	1.44 (0.94 – 2.20)	1.24 (0.38 – 4.12)	1.83 (0.99 – 3.39)	
Viremia, NOS	1.77 (0.97 - 3.23)	N/A	0.91 (0.28 – 2.96)	
Voice disturbance	0.98 (0.59 – 1.61)	0.24 (0.03 – 1.81)	0.98 (0.48 – 1.99)	
Wheezing	1.83 (1.33 – 2.51)*	2.19 (1.08 – 4.46)*	1.60 (0.95 – 2.70)	

Associations with a death outcome are shown using odds ratios with 95% confidence intervals. Results from all three outcomes and sex-stratified associations can be found in Table S4. Transition dates from the pre-Delta to the Delta and Omicron variants were set at 7/3/2021 and 12/17/2021 respectively. All association calculations adjusted for age, sex, race/ethnicity, obesity category, and breakthrough infection status. N/A: no observed outcomes. *: $p < 0.05^+$: $p < 1.1 \times 10^{-4}$ (Bonferroni significance).

Cade *et al.*: Common pulmonary disease and Covid-19 severity Table 5. Joint pulmonary and sleep disorders LASSO regression nonzero coefficients.

Disease or covariate	Death Outcome	Mechanical ventilation or ICU admission Outcome	Inpatient admission Outcome
Age	1.610	0.780	0.826
BMI	0	0.076	0
Sex (1M, 2F)	-0.089	-0.149	-0.042
Breakthrough Infection	0	-0.074	-0.250
Encounters	0.071	0.116	0.296
Days into pandemic (1/1/2020)	-0.442	-0.734	-0.222
Emphysema	N/A	0.071	N/A
Pneumonia	0.039	N/A	0.105
Chronic airway obstruction	0.093	N/A	N/A
Pulmonary congestion and hypostasis	N/A	0.059	0.097
Respiratory failure	0.091	N/A	0.075
Shortness of breath	N/A	N/A	0.141

Nonzero coefficients from adaptive LASSO regressions predicting each outcome using all pulmonary and sleep diseases and covariates are shown. Greater absolute values indicate greater relative importance. The covariates age, sex, BMI, race/ethnicity, breakthrough infection status, total healthcare encounters, and days into the pandemic were forced into all training, testing and combined sample runs. Sex-stratified results are listed in Table S5. N/A: not included in final combined sample run. Asociation odds ratios for individual diseases are presented in Tables 2 and S2.

Cade et al.: Common pulmonary disease and Covid-19 severity

391 Acknowledgments

- 392 The authors wish to thank the MGB Research Patient Data Registry (RPDR) and
- 393 patients for providing health information data.

	_	_	_	_	-	_	_	_	-	_	_		_			_						
ARDS -																						
Asphyxia and hypoxemia -																						
Bacterial infection NOS -																						
Bronchitis -																						
Chronic cirwov obstruction -																						
Chronic all way obstruction -																						
Chronic bronchitis -																						
Cough -																						
Emphysema -																						
Empyema and pneumothorax -																						
Epistavia and priodinorial																						
Influenza -																					Point E	stimate
Lung cancer -																					Differer	nce
Lymphangioleiomyomatosis -																					(0.0
Other diseases of lung -																						
Other dycapace																						-0.5
Pleurisy; pleural effusion -																						
Pneumonia -																						-1.0
Pulmonary collapse; interstitial and compensatory emphysema -																						
Pulmonary congestion and hypostasis -																						-1.5
Pulmonary embolism and infarction acute -																						
Pulmonary hypertension -																						
Respiratory failure -																						
Restless legs syndrome -																						
Shortness of breath -																						
Sleep appea -																						
Solitory pulmonary podulo -																						
Symptoms involving respiratory system and other chest symptoms -																						
Viral infection -																						
Wheezing -																						
	-	•	-	-	'	•	I	-	-	-	-	•	•	-	۱ (-	•	-	-	-		
	õ	Ъ	ure	ab.	lse	ISe	lse	.ab	ISe	.ab	ab	nre	lse	ab.	.ab	õ	ISe	Jre	ab	õ		
	Ē)	Ś	edt	Ē	sea	sea	sea	L L	sea	L L	Ē	edt	sea	Ľ.	°.	Ē)	sea	edt	E L	Ē		
	in	iet	ÖÖ.	dth	Ö	ö	ö	bin	Ö	crit	nni	õ	ö	DUC	tes	ec	ö	ÖÖ.	ate	S		
	0m	al d	Ъ	\geq) L	ia (.e	glo	e.	ato	00	2	IS (ö	SCV	Sp) L	Р	Г	0		
	ŝ	Ш	SS	NO	ctio	oni	ilur	0L	ilur	Ĕ	ell	S	atu	bir	hc	d	sio	SS	atio	es		
	nt;	lor	ğ	outi	Ind	Ę	fa	ler	fa	Ť	0	ğ	lst	glo	Ē	ШO	ten	ğ	Itra	niti		
	ne	~	E	trik	bst	nel	nal	-	ory		Õ	다	nta	UQ	it ly	Õ	ler	an	ur fi	rer		
	ger		gra	dis	0 /	ď	D.		rati		ld l	eal	ner	her	Cen	ode	dуг	SCS	ula	ext		
	naj		dio	ell	va		nic		spi		g	eh	b	ar	ero	ž	a	F	Jer	of		
	ma		arc	0 D	air		JLO		Re			Че	ere	cul	ш.	Ë	snti	er (D	ds		
	Je I		lo	Вē	ji.		Ö						Alto	Sno		\geq	SSE)the	st g	nn		
	arç		ect		ror									orp		\geq	Ш́	0	ш	Ŵ		
	Sch		Ш		Ч									Ö		g				en		
	diś													ea		F-				g		
	ital													\geq		Ċ				-		
	dsu															hrc						
	H															Ē						
																shc						
							_		_							Щ						
							A	dju	iste	ed (CO/	/ar	iat	e								