Here, we report that COVID-19 hospitalization rates follow an exponential relationship with age, doubling for every 16 years of age or equivalently increasing by 4.5% per year of life ($R^2 = 0.98$). This mirrors the well-studied exponential decline of both thymus volume and T-cell production, which halve every 16 years. COVID-19 can therefore be added to the list of other diseases with this property, including those caused by methicillin-resistant *Staphylococcus aureus*, MERS-CoV, West Nile virus, *Streptococcus pneumoniae* and certain cancers, such as chronic myeloid leukaemia and brain cancers.

In addition, the incidence of severe disease and mortality due to COVID-19 are both higher in men, consistent with the degree to which thymic involution (and the decrease in T-cell production with age) is more severe in men compared to women. Since these properties are shared with some non-contagious diseases, we hypothesized that the age dependence does not come from social-mixing patterns, i.e. that the probability of hospitalization given infection rises exponentially, doubling every 16 years. A Bayesian analysis of daily hospitalizations, incorporating contact matrices, found that this relationship holds for every age group except for the under 20s. While older adults have fewer contacts than young adults, our analysis suggests that there is an approximate cancellation between the effects of fewer contacts for the elderly and higher infectiousness due to a higher probability of developing severe disease. Our model fitting suggests under 20s have 49–75% additional immune protection beyond that predicted by strong thymus function alone, consistent with increased juvenile cross-immunity from other viruses. We found no evidence for differences between age groups in susceptibility to infection or infectiousness to others (given disease state), i.e. the only important factor in the age dependence of hospitalization rates is the probability of hospitalization given infection. These findings suggest the existence of a T-cell exhaustion threshold, proportional to thymic output and that clonal expansion of peripheral T-cells does not affect disease risk. The strikingly simple inverse relationship between risk and thymic T-cell output adds to the evidence that thymic involution is an important factor in the decline of the immune system with age and may also be an important clue in understanding disease progression, not just for COVID-19 but other diseases as well.

1. Introduction

Epidemiological patterns in the incidence of a disease can provide insight into the mechanisms of disease progression [1–4]. The degradation of the adaptive immune system with age is already acknowledged to be a major risk factor for both infectious and non-infectious diseases and may play a role in understanding the emerging COVID-19 epidemic. Thymus volume, and the concomitant production of T-cells, decrease exponentially with age with a
2. Results

2.1. COVID-19 hospitalization rates

While data on confirmed cases can be highly variable and largely influenced by testing strategies, the data on hospitalizations, which is the focus of this paper, are relatively more reliable. The incidence of COVID-19 hospitalizations, in a number of countries, consistently doubles with every 16 years of age ($R^2 = 0.98$ for top three countries; figure 1a). Meanwhile, the incidence of all confirmed cases (including mild or asymptomatic) appears roughly constant across adult ages (electronic supplementary material, figure S1). One explanation that is consistent with the data is that exposure is approximately uniform for adult age groups and that after exposure, the probability of becoming hospitalized is proportional to $e^{0.044t}$, where $t$ is age. We will address the age dependence of exposure in more detail by accounting for assortative social mixing as well as a range of additional age-dependent factors in our Bayesian model (see below).

There is a sex bias in COVID-19 risk, which increases with disease severity (figure 1c). This is similar to other diseases, including cancer, where men have 1.33 times the risk of hospitalization and 1.89 times the risk of death [26,29]. The sex bias in COVID-19 is remarkably similar to a factor of 1.35 ± 0.4 for hospitalization incidence and 1.9 ± 0.4 for mortality (mean ± s.d.; figure 1c). The slope of the logarithm of the COVID-19 mortality curve is over twice that of the hospitalization curve, corresponding to an exponential with rate 0.109 ± 0.005 yr$^{-1}$ (figure 1b). Another way of thinking about the sex bias would be to say that for both hospital incidence and mortality, men are effectively approximately 6 years older than women in terms of risk. Other risk factors such as BMI can also be viewed similarly to give an individualized effective ‘Covid age’ [30]. The increase in mortality with age may also be explained by co-morbidities which increase with age, such as cardiovascular disease, which rises exponentially [31] with a rate of 0.071 ± 0.003 yr$^{-1}$. Since 0.071 ± 0.044 = 0.115 ± 0.109, a simple model where the risk of COVID-19 mortality is proportional to the risk of cardiovascular disease and inversely proportional to T-cell production would have the correct age dependence. This would suggest that cardiovascular disease is a risk factor just for the stages in between hospitalization and death.

2.2. Bayesian model

Similar to other diseases, COVID-19 hospitalization risk is relatively high for very young children (e.g. 0.6 cases per 100 000 for ages 0–4 versus 0.2 cases per 100 000 for ages 5–17 in USA; figure 1a). Additionally, older children have a risk lower than expected based on the exponential increase with age we have identified (figure 1a). This is similar to MRSA and S. pneumoniae infection, but not West Nile virus infection or cancers with similar exponential behaviour [4]. Potential factors underlying the apparent low risk in juveniles include age dependence in (i) exposure (e.g. due to heterogeneous social mixing among age groups), (ii) disease progression, (iii) infection given exposure and/or, (iv) infectiousness to others. Throughout this paper, we use the term ‘severe infection’ synonymously with hospitalization and we categorize all infections as either mild or severe. In a preliminary analysis, we first incorporated contact matrices into a simple analytically tractable
susceptible–infected–removed (SIR) model to predict the steady state of the age distribution of hospitalizations in France, with the assumption that the probability of severe disease given infection is proportional to $e^{0.044t}$ (see electronic supplementary material, figure S3). This model suggested that age differences in social mixing could, in part, account for the relatively low hospitalization of non-adults (electronic supplementary material, figure S3). However, the other possible factors in low juvenile COVID-19 hospitalization were not considered in this preliminary analysis.

To incorporate all relevant factors, and to rigorously test the hypothesis that the probability of hospitalization given infection rises with age at the same rate as thymic involution, we conducted a more detailed analysis of age dependence based on daily hospitalization, recovery and death data. We focused on the single country France, for which an unusually comprehensive age distributed dataset is available [20]. All cases in the dataset are either biologically confirmed or present with a computed tomographic image highly suggestive of SARS-CoV-2 infection, and the dataset includes corrections for reporting delays [20]. We formulated an age-structured Bayesian SIR model of infection, partitioning the force of infection into that arising from contacts with mild and severely infected individuals, weighted by age-dependent contact matrices, as well as contact-independent (environmental) transmission. The model-fitting exercise focused on inferring a posterior parameter distribution for the probability of severe disease given infection for each age cohort. In addition, posterior distributions were inferred for a range of secondary parameters (electronic supplementary material, table S2, parameters of the Bayesian analysis), including age-dependent transmissibility and susceptibility.

Our results reiterate that the probability of severe disease given infection increases exponentially with age, at a rate that is remarkably well matched by the rate of thymus decline for all age groups above 20 years (electronic supplementary material, figure S4; all adult age groups have 95% credible intervals including the rate of thymus decline). In order to investigate the nature of juvenile deviation from this exponential relationship, we reformulated the analysis to allow deviations from an exponential increase (for the probability of severe disease given infection) for each age cohort (figure 2). The posterior parameter distribution for the exponential rate was found to match the rate of thymus degradation (95% CI 0.043–0.053 yr$^{-1}$; figure 2b). Only the juvenile age cohort was found to significantly deviate from the exponential response (figure 2c), showing a level of additional protection to severe COVID-19 of between 49 and 75% (electronic supplementary material, table S1). Our sensitivity analysis allowed—within each age cohort—for deviation from uniform the probability of infection given exposure, and, deviation from uniform infectiousness of infected individuals. For both of these, we found that none of the age cohorts deviated significantly (in all cases 95% credible intervals included zero deviation; electronic supplementary material, figure S5), allowing us to discount these potentially confounding factors.

The low susceptibility to severe disease given infection in non-adults may be due to cross-protection from other coronaviruses [8,32–34], or even non-specific protection from other respiratory viruses [35], which occur more frequently in non-adults compared to adults [36]. Our estimate of 49–75% protection ties in with a study which found SARS-CoV-2 reactive antibodies in approximately 60% of unexposed individuals aged 6–16 and only 6% in adults [37]. There is also evidence of unexposed individuals having SARS-CoV-2 reactive CD4+ T-cells [38]. Another possible explanation for the low risk in non-adults might come from some intrinsic feature of the immune system.
For example, we speculate that the cause may be related to the high risk of T-lymphoblastic leukaemia for ages 5–20 (see [4]; electronic supplementary material, figure S5).

3. Discussion

Although we have demonstrated a clear relationship between the probability of severe disease and age, it is possible that the relationship is due, in part, to alternative physical processes other than T-cell production, such as age-related changes in the bone marrow, spleen or lymph nodes. Bone marrow also shrinks with age, but at a rate that is substantially slower than the thymus [39]. Further experiments are needed to determine the degree of causation between T-cell production and disease risk, for example, by measuring TRECs and performing a prospective study or quantifying the increase in risk in thymectomized individuals [40]. Furthermore, a mechanism for why the probability of hospitalization is inversely proportional to T-cell production is currently lacking. One possible model features stochastic fluctuations in the number of infected cells and an immune escape threshold which is proportional to T-cell production [4]. This model has the added benefit that it can also explain most of the other (non-exponential) relationships between risk and age seen in various cancer types [4].

Chronic myeloid leukaemia (CML) is a type of cancer with an age dependence remarkably similar to COVID-19. In both diseases, the risk of hospitalization rises exponentially, inversely proportional to T-cell production [4], with sex bias ratios of 1.35 ± 0.4 for COVID-19 and 1.35 ± 0.3 for CML. The mortality risk profiles are also similar (exponential rates: 0.109 ± 0.005 yr⁻¹ for COVID-19 and 0.103 ± 0.007 yr⁻¹ for CML; sex bias ratios: 1.9 ± 0.4 for COVID-19 and 1.8 ± 0.6 for CML; electronic supplementary material, figure S2). CML is characterized by a single genomic feature, a chromosomal translocation known as the Philadelphia chromosome. This suggests that the probabilities of Philadelphia chromosome formation and COVID-19 infection are approximately age independent, but that the probabilities of subsequent hospitalization are T-cell dependent. A good candidate for a potential mechanism involves the phenomenon that increased antigenic load can lead to T-cell exhaustion, characterized by low effector function and clone-specific depletion [41]. T-cell exhaustion is a factor in both cancer and infectious diseases, including COVID-19 [42,43], where it has even been shown to be a predictor of mortality [44]. As T-cell production decreases with age, this may lead to an increase in the probability of T-cell exhaustion. In support of this hypothesis, low precursor T-cell numbers have been shown to lead to T-cell exhaustion and disease progression in a mouse cancer model [45]. More specifically, we predict a step in disease progression with a probability exactly inversely proportional to the number of precursor T-cells.

When looking at sex biases for COVID-19 hospitalization and mortality (figure 1c), we found factors of 1.35 and 1.9, respectively. We can speculate that since 1.35² ≈ 1.9, this might be an indication that among the steps of disease progression, there could be two T-cell-dependent steps, one pre-hospitalization and one post-hospitalization. This would imply that the risk of death would involve two factors of e⁰.⁰⁴⁴ and therefore mortality would increase at least as fast as e⁰.⁰⁸⁸. The log-slope of the mortality curve being 0.109 ± 0.005 yr⁻¹ is consistent with this hypothesis. One feature of post-hospitalization disease progression is an IL-6 driven cytokine storm [46], which has been related to T-cell dysfunction in a mouse model [47]. These T-cell dysfunction-related cytokine storms were attenuated by nicotinamide adenine dinucleotide precursors and blocking of TNFα signalling.

4. Conclusion

Here, we have shown that the risk of COVID-19 hospitalization rises exponentially with age, inversely proportional to T-cell production, in a similar way to several other diseases. Consistently, the sex bias in disease risk also fits this trend. These features suggest that the risk of hospitalization is related to an immune deficiency, rather than an immunopathology. By contrast, long-COVID follows patterns similar to autoimmune diseases, with middle-aged women having the highest risk [48]. In addition, we found that the under-20 age group benefits from additional protection from severe disease by a factor similar to the prevalence of SARS-CoV-2 cross-reactive antibodies. Our mathematical model suggests that the age dependence of hospitalization...
rates does not arise from differences in social-mixing patterns, but rather from the probability of hospitalization given infection. The model could be easily extended to assess which age groups and socioeconomic groups would be most valuable to vaccinate and therefore to optimize vaccination strategies.

These findings add to the growing evidence that thymic involu- tion is a major component of immunosenescence and that restoring thymus function may be an effective preventative measure for many common diseases. Additionally, our understanding of host–pathogen dynamics is not complete. There is currently no detailed mechanistic model to explain why the probability of hospitalization would be proportional to the reciprocal of thymic T-cell production, for COVID-19 or any other disease. We hope that these findings will be an important clue in understanding the precise mechanisms involved in disease progression.

Data accessibility. All data and code are available in the main text or the electronic supplementary material, or at https://github.com/samindo/COVID.

Authors’ contributions. S.P. developed the initial idea, collected data, did the statistical analysis and drafted the manuscript. N.C. contributed to the interpretation of the data and results and to critically revising the manuscript. R.D. collected data, developed Bayesian modelling, contributed to the interpretation of the data and results and to critically revising the manuscript. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

Competing interests. We declare we have no competing interests.

Funding. We acknowledge financial support from Wellcome Trust (grant no. ID 211944/Z/18/Z).

Acknowledgements. We would like to thank Mark Colles, Eamon Gaffney, Jon Chapman, Philip Maini and Marc-Andre Rousseau for comments. In particular we would like to thank Theo Newman for help with an early draft and for the idea to think of risk in terms of effective ages.

References
processes. PLoS ONE 8, e62390. (doi:10.1371/journal.pone.0062390)

29. Cook MB, McGlynn KA, Devesa SS, Freedman ND, Anderson WF. 2011 Sex disparities in cancer mortality and survival. Cancer Epidemiol. Biomark. Prev. 20, 1629–1637. (doi:10.1158/1055-9965.EPI-11-0246)

30. Coggon D, Croft P, Cullinan P, Williams A. 2020 Assessment of workers personal vulnerability to covid-19 using covid-age. medRxiv 2020.05.21.20108969. (doi:10.1101/2020.05.21.20108969)

31. Waters A-M, Trinh L, Chau T, Bourchier M, Moon L. 2013 Latest statistics on cardiovascular disease in Australia. Clin. Exp. Pharmacol. Physiol. 40, 347–356. (doi:10.1111/1440-1681.12079)

32. Nickbakhsh S, Ho A, Marques DFP, McMenamin J, Gunson RN, Murcia PR. 2020 Epidemiology of seasonal coronaviruses: establishing the context for the emergence of coronavirus disease 2019. J. Infect. Dis. 222, 17–25. (doi:10.1093/infdis/jiaa185)

33. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. 2020 Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science 368, 860–868. (doi:10.1126/science.abb5793)

34. Cohen SA, Kellogg C, Equils O. 2020 Neutralizing and cross-reacting antibodies: implications for immunotherapy and SARS-CoV-2 vaccine development. Hum. Vaccines Immunother. 17, 84–87. (doi:10.1080/21645515.2020.1787074)

35. Cowling BJ et al. 2012 Increased risk of noninfluenza respiratory virus infections associated with receipt of inactivated influenza vaccine. Clin. Infect. Dis. 54, 1778–1783. (doi:10.1093/cid/cis037)

36. Tsagarakis NJ, Sideri A, Makridis P, Triantafyllou A, Stamoulakatou A, Papadogeorgaki E. 2018 Age-related prevalence of common upper respiratory pathogens, based on the application of the FilmArray Respiratory panel in a tertiary hospital in Greece. Medicine 97, e10903. (doi:10.1097/MD.0000000000010903)

37. Ng KW et al. 2020 Pre-existing and de novo humoral immunity to SARS-CoV-2 in humans. bioRxiv 2020.05.14.095414. (doi:10.1101/2020.05.14.095414)

38. Grifoni A et al. 2020 Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. Cell 181, 1489–1501.e15. (doi:10.1016/j.cell.2020.05.015)

39. Blank CU et al. 2019 Defining ‘T cell exhaustion’. Nat. Rev. Immunol. 19, 665–674. (doi:10.1038/s41577-019-0221-9)

40. Zheng M, Gao Y, Wang G, Song G, Liu S, Sun D, Xu Y, Tian Z. 2020 Functional exhaustion of antiviral lymphocytes in COVID-19 patients. Cell. Mol. Immunol. 17, 533–535. (doi:10.1038/s41423-020-0402-2)

41. De Biasi S et al. 2020 Marked T cell activation, senescence, exhaustion and skewing towards TH17 in patients with COVID-19 pneumonia. Nat. Commun. 11, 3434. (doi:10.1038/s41467-020-17292-4)

42. Malandro N et al. 2016 Clonal abundance of tumor-specific CD4+ T cells potentiates efficacy and alters susceptibility to exhaustion. Immunity 44, 179–193. (doi:10.1016/j.immuni.2015.12.018)

43. Moore JB, June CH. 2020 Cytokine release syndrome in severe COVID-19. Science 368, 473–474. (doi:10.1126/science.aab8925)

44. Desdín-Micó G et al. 2020 T cells with dysfunctional mitochondria induce multimorbidity and premature senescence. Science 368, 1371–1376. (doi:10.1126/science.aax0860)

45. Sudre CH et al. 2020 Attributes and predictors of long-COVID: analysis of COVID cases and their symptoms collected by the Covid Symptoms Study App. medRxiv 2020.10.19.20214494. ( doi:10.1101/2020.10.19.20214494)