Data, assumptions, methods and detailed results for the article 'A sensible COVID-19 exit strategy for the UK'

Nicholas Lewis 22 April 2020

Population data

I use the latest official data, for mid-2018, regarding the age-structure of the UK population.[1]

Infection fatality rates and attack rates

I use the Verity et al.(2020)[2] IFR by age-group estimates, modified for 70+ year olds as discussed below [Methods]. Ferguson et al. also used Verity et al. IFR estimates, however they modified them to reflect an assumed non-uniform attack rate by age. I deduce that Ferguson et al. (2020)[3] assumed that the final proportion of the population attacked (infected) by COVID-19 would increase with age, with age-groups older than 50-59 having a significantly higher attack rate than younger age-groups [Methods].

The Ferguson et al. assumption of attack rate increasing with age, for which they cite no supporting evidence, appears to be unjustified. Moreover, the 'Do nothing' assumption in Ferguson et al. that 81% of the UK population would ultimately be infected appears to be inconsistent with its baseline assumption for the COVID-19 reproduction ratio, of R_0=2.4. Those two parameters are mathematically related.[4] If R_0=2.4, one would expect between 87% and 88% of the population to have been infected.

The UK attack-rate estimates calculated in a more recent study by the Imperial College COVID-19 response Team, Walker et al. (2020),[5] *decline* substantially with increasing age, after age 50, reflecting a steep decline in contact patterns with age. Moreover, their attack rate estimates imply that 87% of the UK population would be infected, and so are consistent with the Ferguson et al. baseline R_0=2.4 value. I have accordingly adopted that 87% proportion.

Although I consider the Walker et al. attack rate estimates to be reasonable, I have not adopted them. That is because, even assuming – which I do – a uniform-by-age attack rate, the proportions of deaths in the 80+ age group, and to a lesser extent in the 65+ age group, are substantially lower than indicated by actual deaths in the UK (based on data to 3 April 2020). In the UK and in all other west European countries, at least 50% of total deaths have been in the 80+ age group, and the relative risk of death for people below age 65 has been no more than 0.03 times that for people aged 65+.

In order to approximately match the UK (England and Wales) COVID-19 related deaths by decadal age-group statistics,[6] on the assumption that they reflect attack-rate adjusted IFRs, while using a uniform-with-age attack rate, I found it necessary to scale the Verity et al. IFRs for the 70-79 and 80+ age groups by factors of respectively 1.25 and 2.3 times. After doing so, the ratio of the IFRs for those two age groups became quite close to their ratio as estimated using data from the Diamond Princess cruise ship.[i] I then rescaled all the resulting IFRs to bring the overall IFR for the UK population back to the same level as implied by the unadjusted Verity et al. IFRs, of 0.90%.[ii]

---

[i] The Diamond Princess data (Field Briefing: Diamond Princess COVID-19 Cases, 20 Feb Update. National Institute of Infectious Diseases, Japan https://www.niid.go.jp/niid/en/2019-ncov-e/9417-covid-dp-fe-02.html; update of developments at https://www.mhlw.go.jp/stf/newpage_10989.html) are primarily informative about deaths in the 70-79 and 80+ year age groups, since there was only one case where the age at death (missing in two cases) was known to be under 70 years: https://en.wikipedia.org/wiki/2020_coronavirus_pandemic_on_Diamond_Princess.

[ii] Ferguson et al. similarly scaled the Verity et al. IFR estimates by age-group, albeit with a much less pronounced upscaling for the 80+ age group, and then rescaled, giving a 0.9% overall UK IFR estimate. However they treated the

My resulting estimates of the proportion of COVID-19 related deaths are broadly in line with those implied by decadal age group data from Italy for deaths to early April 2020. They are also closely in line with those implied by the broader age group data published for France.[7] My estimate of the average age at death, being approximately 78 years, is close to that for Italy (78.5 years), and marginally lower than that for the Netherlands (81 years).

It appears that in about 30% of symptomatic cases the standard RT-PCR test for COVID-19 infection gives a negative result when the patient is in fact infected (Ai et al. 2020)[8]. False negative test results will have led to an underestimation of the prevalence of infections both in expatriates evacuated from Wuhan, and in Diamond Princess occupants, and hence biased up both the Verity et al. and the Diamond Princess-based IFR estimates. I therefore reduce, uniformly across age groups, the foregoing IFR estimates to allow for the incidence of false negative infection test results. Conservatively, I apply a reduction of only 15%, half the reduction implied by the Ai et al. findings. Finally, I interpolate IFRs from decadal age groups to pentadal age groups [Methods].

Evidence is emerging that true overall IFRs may be much lower than those implied by the Verity et al. estimates, even after adjusting for false negative RT-PCR test results. For instance, test results from Los Angeles county[9] suggest a best estimate of 331,000 adults had been infected by COVID-19 and had recovered by 10/11 April. Since it takes ~10 days post disease onset for antibodies to be detectable[10], almost all these people would all have started the disease by 31 March (having been infected ~5 days earlier). Since 31 March is the latest likely date of their disease onset, and the mean time from disease onset to death is ~18 days, most of the deaths in that group would have occurred by 19 April, the latest date to which LA county deaths had been updated at the time of writing. Of the 600 total deaths[11], almost all would have been of adults. If one assumed, probably optimistically, that 90% of the deaths of the estimated 331,000 people whose disease started by 31 March had occurred by 19 April, the implied IFR would be (600/0.9)/331,000 = 0.20%. If only 75% of the deaths had occurred, perhaps a better assumption, the IFR would be 0.24%. However, the LA population has a lower proportion of older people than in the UK. If I adjust the implied LA IFR to the UK population structure, using the Verity et al 2020 (Imperial College team) estimate of IFR by age group, it would become 0.25% on the 90% of deaths by 19 April basis, or 0.3% on the 75% of deaths by 19 April basis.

Variation of relative IFRs with chronic health conditions

There is substantial evidence that, in addition to the IFR increasing strongly with age, it is also far higher where a person suffers from one or more relevant health conditions. The prevalence of such conditions itself increases with age. I use estimates of the prevalence of chronic disorders in Scotland by age group (Barnett et al. 2012)[12], dividing each age-group into sub-groups with zero, one, two, or more than two chronic health conditions. Although Barnett et al. cover other disorders as well as relevant health conditions, it appears to be the most suitable analysis.

I set the IFR for cases where respectively one or more than two chronic disorders are present, relative to that for cases with no chronic disorders, at respectively 4 and 6 and use an interpolated relative IFR of 5 for cases with two chronic disorders. These ratios are applied to all age groups, the available data being inadequate to estimate any age dependency. The values of 4 and 6 were chosen to achieve a reasonable match to available data on the breakdown of deaths by the number of chronic disorders in Italy, that being the only country for which sound such data were readily available.[13] The Italian data

---

scaling as relating to higher attack rates for older age-groups, which is totally at variance with the Walker et al. estimates.

are based on chronic diseases listed in patients' hospital charts. An analysis of the Italian data that takes account of the different, generally lower, prevalence of chronic disorders in Italy than per the Scotland-based data I use implies that considerably higher relative IFR ratios for chronic disorders than I adopt would be justified: a ratio of approximately 10, averaged over different numbers of conditions.[14]

Much of the available data on the effects of chronic disorders on COVID-19 fatality risk is from China, but it is unclear how applicable analyses based on it are to the UK. The age-breakdown of COVID-19 deaths in China is significantly different from that in Europe, and a large proportion of cases appear to have been unreported (Verity et al.), which may have substantially skewed the analyses.

Data for New York City, which is available on a 2-dimensional basis by broad age groups and the presence or not of any underlying health condition, indicates that there an even lower proportion of deaths occur when no relevant health conditions are present than that implied by the scaling factors that I use, suggesting that higher scaling factors might be appropriate. However, the age-breakdown of COVID-19 deaths in New York City is significantly different from that in European countries.

Data from the Netherlands[15] indicate that a higher proportion of total deaths have occurred in cases where no relevant health conditions were present than that implied by the scaling factors that I use, suggesting that smaller scaling factors might be appropriate. However, the Netherlands data for deaths up to age 70[16] do not show a significantly higher proportion in cases where no relevant health conditions were present than the proportion that I derive.[iii]

Data published for the UK are not useful, as they only relate to cases where no underlying health condition was mentioned on the death certificate. Only significant conditions that are considered to have contributed to a death are listed. In most cases where a patient was known to have had a relevant health condition, it seems unlikely that it would be known to be a significant contributor to the death.

Proportion of asymptomatic cases

Ferguson et al. assumed that only one-third of COVID-19 infections were asymptomatic, however this figure appears unrealistically low in the light of the evidence now available.

The Ferguson et al. assumption was itself lower than the proportion suggested by data from China and repatriation flights from Wuhan, used in Verity et al. Ferguson et al. stated using that these data suggested that 40-50% of infections were not identified as cases. However, the ratio of the Verity et al. overall IFR estimate of 0.657% to their fully-adjusted case fatality rate (CFR) estimate of 1.38% appears to imply that slightly over 50% of infected cases were asymptomatic.

Consistent with this, complete testing of occupants of the Diamond Princess indicated that between 50% and 60% of confirmed infected cases were asymptomatic at the time of testing, with the proportion remaining asymptomatic falling to slightly under 50% a month later.[iv] Complete testing of

---

[iii] Data published for France (https://www.santepubliquefrance.fr/content/download/243797/2565951) also suggest that a higher proportion of total deaths have occurred in cases where no relevant health conditions were present. However, the French data show death occurring in only 8% of intensive care patients, a far lower proportion than usual, which suggests that the vast majority of deaths of patients in intensive care, a high proportion of whom are likely to have relevant health conditions, are yet to be reported.

[iv] The average age of Diamond Princess cruise ship occupants was relatively high, and the proportion of symptomatic infections appears to increase with age (Wu20), so the proportion of Diamond Princess infections that were asymptomatic may be a biased low estimator for the population as a whole. However, the Diamond Princess data itself, leaving aside age groups with tiny sample sizes (<19 and >90 years) and very old suggested that the proportion of symptomatic infections declined with age.

the population of Vo, Italy, indicated that 50-75% of infections were asymptomatic.[17] Random testing of the Icelandic population indicated that about 50% are asymptomatic.[18] Data from China suggests that some 78% of infections are asymptomatic.[19]

In the light of the foregoing evidence, I adopt the assumption that 50% of COVID-19 infections are asymptomatic; the proportion appears unlikely to be lower than that. Similarly to Ferguson20, I assume that the asymptomatic proportion is not age dependent, since the data are inadequate to estimate any age dependency.

Hospitalisation and critical care (ICU) rates

I adopt the Ferguson et al. estimates of the proportion of symptomatic cases that are hospitalised, after removing the attack-rate adjustments that they had applied [Methods], and apply them using instead my assumptions as to attack rates and the proportion of cases that are symptomatic. I also adopt the Ferguson et al. estimates of the proportion of hospitalised cases requiring critical care/ICU.

Results

The estimates that I derive based on the data and assumptions described above are set out in Table 1. The calculations involve straightforward arithmetic.

Table 1: Projected deaths and ICU usage by age group and relevant health conditions status

Age group (years)	UK mid-2018 population (thousands)	Estimated no. with no relevant health condition (thousands)	Overall Infection Fatality Ratio (IFR)	Estimated IFR if no relevant health conditions	Deaths if 87% of population infected (1)	Deaths if groups with relevant health conditions not infected	ICU beds occupied if 87% of population infected	ICU beds if groups with relevant health conditions not infected
0–9	8,053	7,524	0.001%	0.001%	0	0	0	0
10–19	7,528	6,202	0.004%	0.002%	0	0	0	0
20–29	8,712	6,855	0.02%	0.01%	1,000	1,000	200	100
30–39	8,836	6,099	0.05%	0.02%	4,000	1,000	500	200
40–49	8,501	4,958	0.09%	0.04%	7,000	2,000	900	300
50–59	8,968	3,918	0.33%	0.10%	26,000	3,000	3,700	600
60–69	7,070	1,849	1.04%	0.26%	64,000	4,000	9,100	800
70–79	5,487	777	2.84%	0.60%	136,000	4,000	15,200	600
80+	3,282	246	9.94%	1.9%	284,000	4,000	17,000	400
Total	66,436	38,428	0.90%	0.19%	522,000	19,000	46,600	3,000

Notes

(1) For each age-group, restricted where relevant to people with no relevant health conditions, the number of fatalities is the product of the estimated IFR and 87% of the population in that group.

(2) For each age-group, restricted where relevant to people with no relevant health conditions, ICU bed usage is based on the assumed proportions of symptomatic infections requiring hospitalisation and, of those, also requiring critical care, in Table 1 of Ferguson et al., and the Ferguson et al. assumption that the average

length of stay in an ICU bed is 10 days. Appropriate adjustments are made for the different assumptions regarding attack rates and the symptomatic proportion of infections, which are discussed in the text.

(3) Death and ICU bed numbers are rounded to the nearest 1,000 and 100 respectively.

After the aforementioned adjustments to the Verity et al. IFR estimates, the overall IFR for the UK population is 0.90%. When applied in conjunction with the 87% overall attack rate that is consistent with the Ferguson et al. baseline assumption of R_0=2.4, on a uniform-with-age basis, this results in 522,000 projected deaths. That is closely in line with the 510,000 deaths projected in Ferguson et al. in the 'Do nothing' case, despite slightly different assumptions.[v]

Methods

Treatment of age group data

Much of the data used are broken down by, or can be aggregated into, decadal age groups, often with a single group for 80+ years age. Where data for 80+ year olds are broken down into separate age groups, they are combined. Some data are for broader age groups with breakpoints that are a multiple of five years. To enable use of and/or comparisons with such data, decadal age group data is interpolated into pentadal age groups, the oldest group being 85+ years.

The decadal age group IFR estimates change increasingly rapidly with age, but transforming them by raising to the power 0.05 linearises the relationships (resulting R^2 = 0.995). The younger and older pentadal age group estimates corresponding to each decadal age group are formed by adjusting the transformed IFR for that age group downwards or upwards, respectively, by 2.5 times the annualised difference between the transformed IFRs for the adjacent younger and older decadal age groups. For the youngest and oldest decadal age groups, the annualised difference between their IFRs and those for the adjacent age group are used for these adjustments. After linear interpolation, the IFRs are retransformed. The same method and power is used to interpolate the Ferguson et al. ICU usage assumptions for age groups 40+; below that age these are constant. The same method is also applied to interpolate hospitalisation rates to pentadal age groups, save that a power of 0.333 (cube root) is used to transform the data, this being found to make its age-variation most linear (resulting R^2 = 0.989). When presenting results, calculations based on 5-year age groups are aggregated into decadal age groups.

Hospitalisation rates

I adopt the Ferguson et al. Table 1 attack-rate adjusted estimated proportions of symptomatic cases requiring hospitalisation, which have been scaled to match expected rates for the oldest age group in a UK/US context. I convert their estimates to infection hospitalisation rates (IHR) by multiplying by the proportion of cases that Ferguson et al. assumes to be symptomatic. I then remove the estimated Ferguson attack-rate adjustments. To do so, I estimate the relative attack rate by age-group from the ratios of the Ferguson et al. attack-rate adjusted IFRs to the unadjusted Verity et al. IFRs on which they are based; no other scaling was applied to IFRs in Ferguson20. For the 50-59 age group, the Ferguson et al. IFR is identical to the Verity et al. IFR[20]. I can therefore estimate Ferguson20's assumed attack rates for other age-groups relative to the 50-59 age-group. Since the rounded IFR figures available for under-50 age-groups preclude accurate estimation of separate relative attack-rates for each, I have used the same value for all of them, calculated so that the relative attack-rate for the total population is exactly one.[21]

[v] Both these projections assume that the UK healthcare system is not overwhelmed by COVID-19 cases.

References and Notes

[1] Table MYE2 of https://www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland/mid20182019laboundaries/ukmidyearestimates20182019ladcodes.xls

[2] Verity R, Okell LC, Dorigatti I, et al. Estimates of the severity of COVID-19 disease. medRxiv 13 March 2020; https://www.medrxiv.org/content/10.1101/2020.03.09.20033357v1.

[3] Neil M Ferguson et al., Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand, Imperial College COVID-19 Response Team Report 9, 16 March 2020, https://spiral.imperial.ac.uk:8443/handle/10044/1/77482

[4] Proportion not infected = exp(–R0 x Proportion infected)

[5] Patrick GT Walker, Charles Whittaker, Oliver Watson et al. The Global Impact of COVID-19 and Strategies for Mitigation and Suppression, Imperial College London (2020). UK attack rates digitised from Figure 2.A. http://web.archive.org/web/20200329181205/https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-Global-Impact-26-03-2020.pdf

[6] Weekly provisional figures on deaths registered in England and Wales, Office for National Statistics https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/weeklyprovisionalfiguresondeathsregisteredinenglandandwales Accessed 15April 2020.

[7] COVID-19: Point épidémiologique hebdomadaire du 09 avril 2020. Santé publique France. Based on reported data at 7 April 2020.Accessed 15April2020. https://www.santepubliquefrance.fr/content/download/243797/2565951

[8] Tao Ai, Zhenlu Yang, Hongyan Hou, Chenao Zhan, Chong Chen, Wenzhi Lv, Qian Tao, Ziyong Sun, Liming Xia. Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases**.** *Radiology*, 2020; 200642 DOI: 10.1148/radiol.2020200642 https://www.sciencedaily.com/releases/2020/02/200226151951.htm

[9] http://www.publichealth.lacounty.gov/phcommon/public/media/mediapubhpdetail.cfm?prid=2328

[10] https://www.medrxiv.org/content/10.1101/2020.03.30.20047365v2

[11] https://public.tableau.com/views/COVID-19PublicDashboard/Covid-19Public: as at 21 April 2020

[12] Barnett, Karen, et al. "Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study." *The Lancet* 380.9836 (2012): 37-43. https://www.sciencedirect.com/science/article/pii/S0140673612602402 Proportions by age group of cases with 0, 1, 2 or >2 chronic disorders digitised from Figure 1.

[13] COVID-19 Surveillance Group (2020). Characteristics of COVID-19 patients dying in Italy Report based on available data on April 13, 2020. Rome, Italy, Istituto Superiore di Sanita. Accessed 15 April 2020. https://www.epicentro.iss.it/en/coronavirus/bollettino/Report-COVID-2019_13_april_2020.pdf

[14] Gjerstad, Steven, and Andrea Molle. "Comorbidity Factors Influence COVID-19 Mortality Much More than Age." (2020). https://blogs.chapman.edu/wp-content/uploads/sites/46/2020/03/Comorbidity_Italy_2020_04_01.pdf

[15] Rijksinstituut voor Volksgezondheid en Milieu – RIVM. Epidemiologische situatie COVID-19 in Nederland. Published 2020-04-04. Available from: https://www.rivm.nl/documenten/epidemiologische-situatie-covid-19-in-nederland-4-april-2020https://www.rivm.nl/documenten/epidemiologische-situatie-covid-19-in-nederland-4-april-2020. Accessed 10 April 2020.

[16] Rijksinstituut voor Volksgezondheid en Milieu – RIVM. Epidemiologische situatie COVID-19 in Nederland. Published 2020-04-04. https://www.rivm.nl/documenten/epidemiologische-situatie-covid-19-in-nederland-11-april-2020https://www.rivm.nl/documenten/epidemiologische-situatie-covid-19-in-nederland-11-april-2020. Accessed 15 April 2020.

[17] https://corrierefiorentino.corriere.it/firenze/notizie/cronaca/20_marzo_15/dobbiamo-cambiare-rotta-ef23a500-669a-11ea-a40a-86d505f82a96.shtml?refresh_ce-cp . Accessed 15 April 2020.

[18] https://www.visir.is/g/2020140116d/um-helmingur-greindra-i-skimun-einkennalaus-eda-einkennalitill

[19] Covid-19: four fifths of cases are asymptomatic, China figures indicate *BMJ 2020;369:m1375* doi: https://doi.org/10.1136/bmj.m1375 (Published 02 April 2020)

[20] Verity et al. Table 1.

[21] The calculated common Ferguson et al. relative attack rate of 0.936 for under-50 age groups is consistent, within rounding uncertainty, with the Verity20 IFR to Ferguson20 IFR ratios for all separate under-50 age groups. The estimated Ferguson et al. relative attack rates for age groups 60-69, 70-79 and 80+ years are respectively 1.14, 1.185 and 1.20.