Imperial College UK COVID-19 numbers don't seem to add up

Introduction and summary

A study published two weeks ago by the COVID-19 Response Team from Imperial College (Ferguson20[1]) appears to be largely responsible for driving UK government policy actions. The study is not peer reviewed; indeed, it seems not to have been externally reviewed at all. Moreover, the computer code used to produce the estimates in the study – which on Ferguson's own admission is old, unverified and documented inadequately, if at all – has still not been published. That, in my view, shows a worrying approach to a matter of vital public concern.

However, even in the absence of computer code, it should be possible to check some of the key estimates in Ferguson20 – in particular those based on a 'Do nothing' scenario – and hence establish if there are any obvious problems with the study. Since there is no evidence that anyone in government has performed such a verification exercise, I decided to do so. The results suggest either that Ferguson20 has misstated or omitted key assumptions, or that their model does not correctly derive the numbers of people infected, hospitalised and/or dying from the COVID-19 coronavirus.

I use the relevant input assumptions for Ferguson20 (without expressing any opinion on them), along with UK population data, to calculate fatality and hospitalisation figures. The resulting calculated fatality and hospitalisation figures can be expressed as percentages of total infections and compared with the equivalent Ferguson20 values. I also use estimates given in another paper (Verity20[2]) by the same team at Imperial, from which the Ferguson20 assumptions that I use were largely derived, in order to estimate the age-dependency that Ferguson20 assumes for the "attack rate" (the proportion of the population that contracts COVID-19) and then cross-check the Ferguson20 hospitalisation rates. Details of my calculations are set out in the Methods section appended to this article.

Infection and fatality rates and numbers

As Table A shows, the Ferguson20 estimate of 81% of the total population being infected in a 'Do nothing' scenario corresponds to nearly 54 million people in the UK contracting COVID-19 over the course of the epidemic. On that basis, Ferguson20 estimates that 510,000 people would die,[3] implying an infection fatality ratio (IFR) of 0.948%. That is consistent with the rounded figure Ferguson20 gives[4] of 0.9% (the authors ignore the potential negative effect on mortality of the health system being overwhelmed).

However, using Ferguson20's stated assumptions, I calculate a 30% higher death toll of approximately 660,000 people (Column A7), implying an IFR of 1.23%.

In order to cross-check the Ferguson20 hospitalisation rates I need the relative attack rate by age-group used, however Ferguson20 does not state this. I therefore estimate the relative attack rate by age-group from the ratios of the Ferguson20 attack-rate adjusted IFRs to the unadjusted Verity20 IFRs on which they are based. I then use those estimated Ferguson20 relative attack rates (the accuracy of which is validated by comparing the figures in Column A8 with those in Column A7) in my hospitalisation rate analysis.

Table A: Infection Fatality estimates derived from Ferguson20 assumptions and UK population data compared with those stated in Ferguson20.

A1	A2	A3	A4	A5	A6	A7	A8
Age group (years)	UK mid-2018 population	Ferguson20 attack-rate adjusted Infection Fatality Ratio (IFR)	Verity20 estimated IFR, as used by Ferguson 20	Estimated relative attack rate used in Ferguson 20	Implied no. of people infected if 81% of overall population infected (1)	Ferguson20 IFR no. of fatalities if 81% of overall population infected (2)	*Attack-rate adjusted Verity20 IFR: fatalities if 81% of population infected (3)*
0–9	8,052,552	0.002%	0.0016%	0.936	6,105,592	130	*98*
10–19	7,528,144	0.006%	0.007%	0.936	5,707,976	366	*400*
20–29	8,711,750	0.03%	0.031%	0.936	6,605,408	2,117	*2,048*
30–39	8,835,591	0.08%	0.084%	0.936	6,699,306	5,725	*5,627*
40–49	8,500,792	0.15%	0.16%	0.936	6,445,455	10,328	*10,313*
50–59	8,968,055	0.60%	0.60%	1.000	7,264,125	43,585	*43,585*
60–69	7,069,544	2.2%	1.9%	1.140	6,528,017	125,979	*124,032*
70–79	5,487,167	5.1%	4.3%	1.185	5,266,857	226,675	*226,475*
80+	3,281,955	9.3%	7.8%	1.200	3,190,060	247,230	*248,825*
Total	66,435,550			1.000	53,812,796	662,136	*661,402*
As % of infections						**1.23%**	
Per Ferguson20						***0.9%***	

Notes

(1) For each age-group, the implied number of people infected is the product of 81% of the population in that age group (Column A2) and the estimated relative attack rate (Column A5).

(2) For each age-group, the number of fatalities in Column A7 is the product of the Ferguson20 attack-rate adjusted IFR (Column A4) and 81% of the population (Column A).

(3) For each age-group, the number of fatalities in Column A8 is the product of the Verity20 estimated IFR (Column A4) and the implied number of people infected (Column A6). These close match between these fatality numbers and those in Column A7 validates the estimated relative attack rates in Column A5.

Hospitalisation rates and numbers

As Table B shows, the Ferguson20 estimate of 81% of the total population being infected by COVID-19 on a 'Do nothing' scenario and two-thirds of them suffering symptoms corresponds to nearly 36 million people in the UK being infected symptomatically over the course of the epidemic (Column B2). On the basis of Ferguson20's attack-rate-adjusted estimates of symptomatic cases requiring hospitalisation (Column B3), almost 2.8 million people would require hospitalisation (Column B8).

Using Ferguson20's stated assumptions, I thus calculate that 5.17% of infected people would be hospitalised, a 17.5% higher rate than Ferguson20's 4.4%.

Moreover, Ferguson20 states that its hospitalisation rate assumptions are based on figures for estimated proportions of infections that would be hospitalised from Verity20. Applying those

Verity20 estimates (after adjusting from a per-infection basis to a per-symptomatic-infection basis, and for non-uniform attack-rate) (Column B7), suggests that almost 4.5 million infected people would be hospitalised (Column B9), 89% more implied by Ferguson20's figures, and 61% more than the almost 2.8 million people I calculate on the basis of the Ferguson20 stated assumptions.[5] This suggests that Ferguson20 substantially downscaled the Verity20 hospitalisation rates. There is a hint that this may have been done to adjust for the use of input data ultimately derived from a Chinese context into a GB/US one: the paper speaks of scaling the data so that the hospitalisation rates for the 80+ age group matched those expected in a GB/US context.[6] However, since the Verity20 hospitalisation rates were already adjusted to a UK context,[7] it is not clear why the Ferguson20 authors considered it necessary to take this step.

Table B: Hospitalisation estimates derived from Ferguson20 assumptions and Verity20 data compared with those stated in Ferguson20.

B1	B2	B3	B4	B5	B6	B7	B8	B9
Age group (years)	No. of people symptomatic if 81% of population infected, 2/3 symptomatic (attack-rate unadjusted)	Ferguson20 attack-rate adjusted fraction of symptomatic cases hospitalised	Verity20 assumed fraction of infections requiring hospitalisa-tion	Verity20 fraction of infections showing symptoms (0)	Estimated relative attack rate used in Ferguson 20 (1)	Implied Verity20 attack-rate adjusted fraction of symptomatic cases hospitalised	Ferguson20 implied no. hospitalised: 81% of total population infected, 2/3 symptomatic	Verity20 implied no. hospitalised: 81% of total population infected, 2/3 symptomatic
0–9	4,348,378	0.1%	0.00%	61.5%	0.936	0.00%	4,348	-
10–19	4,065,198	0.3%	0.04%	47.3%	0.936	0.08%	12,196	3,218
20–29	4,704,345	1.2%	1.1%	51.7%	0.936	1.99%	56,452	93,754
30–39	4,771,219	3.2%	3.4%	57.5%	0.936	5.53%	152,679	263,931
40–49	4,590,428	4.9%	4.3%	53.3%	0.936	7.55%	224,931	346,443
50–59	4,842,750	10.2%	8.2%	46.2%	1.000	17.77%	493,960	860,395
60–69	3,817,554	16.6%	11.8%	47.5%	1.140	28.32%	633,714	1,081,131
70–79	2,963,070	24.3%	16.6%	50.0%	1.185	39.34%	720,026	1,165,731
80+	1,772,256	27.3%	18.4%	58.2%	1.200	37.93%	483,826	672,257
Total	35,875,197				1.000		2,782,132	4,486,862
As % of infections							**5.2%**	**8.3%**
Per Ferguson20							***4.4%***	

Conclusions

There may be a perfectly good explanation for the substantial apparent discrepancies in the Ferguson20 estimates of hospitalisation and fatality rates from COVID-19 that I have uncovered. Other than actual miscalculations in Ferguson20, they may have misstated or omitted important assumptions, or I may have misunderstood their assumptions or how they apply, or made an error in my calculations. However, until and unless the Imperial College COVID-19 Response Team show that some combination of these possibilities accounts for the apparent discrepancies, all the results of

their study must be treated with care, as they could potentially be significantly in error, even if the assumptions are valid.

Methods Appendix

In the following, I refer to column numbers given in the top rows of Tables A and B.

Ferguson20 gives the estimates of the severity of cases by age-group that they use.[8] The authors say that the percentage of symptomatic cases requiring hospitalisation and the IFR estimates are taken fromVerity20, which were primarily based on Chinese data. However, Ferguson20 states that those estimates were adjusted to account for non-uniformity of attack rate (the proportion of each age-group that becomes infected), with the hospitalisation percentages additionally subjected to an overall scaling to match expected rates for the 80+ age-group in a UK/US context.

Fatalities

Ferguson20 estimates that 81% of the total UK population would become infected in the absence of any policy measures. I therefore apply the Ferguson20 attack-rate adjusted IFRs (Column A3) to 81% of the population figures (Column A2) in order to calculate the number of fatalities they imply (Column A7). Unhelpfully, Ferguson20 does not state the assumed attack rates, which I need later in my analysis. However, for the 50-59 age group, the Ferguson20 IFR is identical to the Verity20 IFR[9] (Columns A3 and A4). I can therefore estimate Ferguson20's assumed attack rates for other age-groups relative to the 50-59 age-group (Columns A5 and B6). 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 attack-rate for the total population is exactly one.[10] Using that value and the estimated relative attack rates, I can then deduce, based on the latest age-structure of the UK population[11] (Column A2), how many people Ferguson20 estimates will become infected in each age-group (Column A6). I can then use those numbers, together with the Verity20 IFRs (Column A4) to estimate how many fatalities are to be expected for each age-group and overall (Column A8) and hence check (by comparing the Column A8 figures with those in Column A7) that my estimates of the relative attack-rates used in Ferguson20 are sufficiently accurate.

Hospitalisations

The pre-attack rate adjustment number of symptomatic infected people in each age-group (Column B2) is derived by scaling its population (Column A2) by the Ferguson20 assumptions that 81% of the population is infected and 2/3 of infections will be symptomatic. I then estimate, by multiplying those numbers of symptomatic infected people by the Ferguson20 attack-rate adjusted fractions of symptomatic cases hospitalised (Column B3), how many people in each age-group will be hospitalised using the Ferguson20 assumptions (Column B8). I can also use the Verity20 Table 3 Percentage of infections hospitalised estimates (Column B4), by first dividing them by estimates of the fraction of infections showing symptoms per the Verity20 results[12] (Column B5) to convert them to percentages of symptomatic infections hospitalised and then multiplying them by the previously estimated relative attack rates (Column B6). Multiplying the resulting attack-rate adjusted hospitalisation rates numbers of symptomatic infected people (Column B7) then gives the implied numbers of people hospitalised using the Verity20 hospitalisation assumptions (Column B9). That estimate will however not reflect the overall age-independent scaling of hospitalisation rates imposed in Ferguson20.

Nicholas Lewis 1 April 2020

---

[1] 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

[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] Their page 7.

[4] Their page 5.

[5] Further investigation suggests that Ferguson20 in fact applied the Verity20 'Percentage of infections hospitalised' estimates to the Ferguson20 attack-rate adjusted number of people infected, and then scaled the resulting estimates by approximately 0.82 overall, uniformly for ages 40+ but with varying scaling for the (much lower hospitalisation rate) younger age-groups. Since the Verity20 hospitalisation rates per infection were derived by converting rates per symptomatic case using Verity20's estimates of the fraction of infections showing symptoms, which are age-varying and all lower than the 2/3 estimate used in Ferguson20, it seems to me inappropriate to use the Verity20 hospitalisation rates per infection in this way.

[6] Ferguson20 Table 1 caption.

[7] Verity20 Table 3 caption.

[8] Their Table 1.

[9] Verity20 Table 1.

[10] The calculated common relative attack-rate of 0.936 is consistent, within rounding uncertainty, with the Verity20 IFR to Ferguson20 IFR ratios for all separate under-50 age groups.

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

[12] For Verity20, the proportion of infections that were symptomatic is estimated by dividing by the ratio of the IFR in the final column of Table 1 of Verity20 to the fully adjusted case fatality rate (CFR) in the penultimate column of that table.