ESTIMATING ETHNIC POPULATIONS BETWEEN CENSUSES

Philip Rees, Paul Norman
School of Geography, University of Leeds, Leeds LS2 9JT, UK
Gabriele Price, Stephanie Webb
Office for Health Improvement and Disparities, Department of Health and Social Care (OHID)r
Corresponding author: p.h.rees@leeds.ac.uk

TOPIC: Background

The UK's first wave of the Covid-19 pandemic in 2020 revealed that higher mortality risks were experienced by ethnic minority health professionals, essential workers and residents. The Office for National Statistics (ONS) created a database of individuals classified by ethnicity in the 2011 census (Nafilyan et al. 2021). These were survived forward and linked to their 2020 records of deaths from Covid-19. The ONS study showed that socio-economic variables accounted for most of the higher risk experienced by some ethnic minority groups. However, because of the highly confidential nature of the data, no summaries were compiled for local or small areas. The NHS compiled records of SARS-CoV-2 cases, hospitalization and deaths for local, health and small areas. The only reliable source for the ethnic composition of the populations of subnational areas was the 2011 Census. As Table 1 show, ethnic minority population in total increased by 40.4% in the projection from 2011 to 2021 and by 35.3% when 2011 and 2021 census populations are compared. Published estimates of Covid-19 cases and death rates by ethnic minority group computed using a 2011 population at risk are likely to show huge errors.

THEORETICAL FOCUS: The research questions and aims

A team at the School of Geography at the University of Leeds has developed and published population projections for ethnic groups based on 2001 and 2011 Census populations. In 2021 a team at the Office of Health Improvement and Disparities (OHID) asked the Leeds team to adapt the ETHPOP projections to account for new information about population change since 2011. OHID required reliable ethnic population estimates for mid-years in the decade. The research questions were: "How could this be done?" and "Were the results reliable?". The aims were to (1) develop methods for adjusting the projected populations to incorporate ONS mid-year population estimates for the years 2011 to 2021, (2) to evaluate the results against the 2021 Census and (3) to convert the estimates from Lower Tier Local Authorities (LTLAs) to health authority geographies, which from mid-2022 are Integrated Commissioning Systems (ICSs).

DATA: ETHPOP projection inputs

Because official estimates of the components of change by ethnicity were not produced for 2011-21 by ONS, effort was devoted to estimating national and subnational estimates of the required rates. The method used for estimating ethnic fertility rates is described Norman et al. (2014). Ethnic mortality rates are estimated using the geographically weighted method in Rees et al (2009). Internal migration rates by ethnicity for 2010-11 were derived from the 2011 Census special migration tables for England with added estimates of migration rates between UK nations using methods described in Lomax (2013) and Lomax et al. (2014). Methods for estimating international migration are explained in the supplementary online document accompanying Lomax et al. (2019).

For England as a whole, ethnic mortality rates did not vary greatly, with life expectancies being within 2 years of the England average. However, they varied significantly across subnational areas with gaps of up to 10 years between the best and worst performing areas (Rees et al. 2009). Ethnic fertility rates were estimated using survey-based own child estimates, census-based child-woman ratios and national age-specific fertility rates. Bangladeshi, Pakistani and Black African total fertility rates (TFRs) were above replacement (2.07 in the UK) in 2010-2011, the base year, while other ethnic group TFRs were lower than replacement. Internal migration rates by ethnicity, for the year prior to census, between local areas were derived from a 2011 census special migration dataset. Ethnic minority groups had lower rates than the White British and Irish majority group but migration directions to areas of lower population density were similar. As estimates of immigration and emigration by ethnicity were not available directly, indirect estimates were produced by reconstructing tables of international migration by country of birth from the official International Passenger Survey (now discontinued) combined with 2011 Census tables linking country of birth and ethnicity at national, regional and local scales (Lomax et al. 2019, SOM1).

METHODS: The ETHPOP projection model

A bi-regional cohort component model was used to project the local ethnic populations by sex and single year of age. Rees, et al (2023) provides details of the ETHPOP model. A bi-regional model was used to reduce the computational load compared with a multiregional model. The latter requires the estimation of origin-destination rates by ethnicity, whereas the bi-regional model requires estimates of total out-migration for LTLAs by ethnicity and out-migration from the rest of the country to supply in-migration to the LTLAs. The model uses single years of age and a time interval of single years. The spatial scale was the set of lower tier local authority areas (LTLAs) for England as defined in the 2011 Census but then converted to 2021 LTLAs. The 18 ethnic groups employed in the 2011 census were reduced to 12 groups in the ETHPOP model, to achieve

harmonization across the UK. Each ethnic group is projected independently, with ethnicity being fixed for the duration of the projection to 2061 but does allow for offspring of mixed ethnicity.

Table 1: Comparison of the population of England by ethnicity from a projection (ETHPOP) and the 2021 Census, adjusted to the post census official mid-year estimate

ETHPOP FORECASTS FOR MID-YEARS							
		Populations	Populations	Changes	Changes	Composition	Composition
		(1000s)	(1000s)	(1,000s)	%	%	%
Ethnic Group Names	Code	2011	2021	2011-21	2011-21	2011	2021
White British & Irish	WBI	42,928	43,163	235	0.5	80.83	75.13
White Other	WHO	2,435	3,820	1,385	56.9	4.58	6.65
Mixed	MIX	1,194	1,638	444	37.2	2.25	2.85
Indian	IND	1,399	1,912	513	36.7	2.63	3.33
Pakistani	PAK	1,114	1,547	433	38.9	2.10	2.69
Bangladeshi	BAN	438	568	130	29.7	0.82	0.99
Chinese	CHI	380	661	281	74.1	0.71	1.15
Other Asian	OAS	821	1,147	326	39.7	1.55	2.00
Black African	BLA	979	1,356	376	38.4	1.84	2.36
Black Caribbean	BLC	592	585	-7.3	-1.2	1.12	1.02
Other Black	OBL	278	321	43	15.3	0.52	0.56
Other Ethnic	OTH	549	734	185	33.7	1.03	1.28
All Groups	SUM	53,107	57,452	4,344	8.2	100.00	100.00
Minority Groups	SUM	10,179	14,288	4,109	40.4	19.17	24.87
CENSUS POPULATIONS ADJUSTED TO MID-YEAR POPULATION ESTIMATES							
		Populations	Populations	Changes	% Changes	Composition	Composition
		(1000s)	1000s	1,000s	%	%	%
Ethnic Group Names	Code	2011	2021	2011-21	2011-21	2011	2021
White British & Irish	WBI	42,851	42,251	-600	-1.4	80.83	74.73
White Other	WHO	2,430	3,584	1,154	47.5	4.58	6.34
Mixed	MIX	1,193	1,667	474	39.8	2.25	2.95
Indian	IND	1,396	1,843	448	32.1	2.63	3.26
Pakistani	PAK	1,112	1,569	457	41.1	2.10	2.78
Bangladeshi	BAN	437	629	193	44.1	0.82	1.11
Chinese	CHI	380	431	51	13.6	0.72	0.76
Other Asian	OAS	819	952	132	16.1	1.55	1.68
Black African	BLA	978	1,468	490	50.1	1.84	2.60
Black Caribbean	BLC	991	620	-371	-37.4	1.87	1.10
Other Black	OBL	278	294	16	5.7	0.52	0.52
Other Ethnic	OTH	548	1,229	680	124.1	1.03	2.17
All Groups	SUM	53,012	56,536	3,524	6.6	100.00	100.00
Minority Groups	SUM	10,561	14,285	3,724	35.3	19.92	25.27

METHODS: ETHPOP projection assumptions
Given the paucity of time series of demographic component rates classified by ethnicity, we assumed constancy over the 2011 to 2021 decade or adopted ONS assumptions in the national projections for all groups. Mortality, in the decade 2001-2011 prior to the projection "jump off" year of 2011, steadily improved. The ONS assumption of an annual improvement of 1.2% (a 100-year average) was adopted. In the 2001-2011 period, fertility increased, because women in their thirties caught up from postponement in their twenties in the previous decade, 1991-2001. However, this tempo effect could not extrapolated into the future and constant ethnic fertility rates were assumed. Internal migration varied with the economic cycle in 2001-2011 (Lomax et al. 2013) but no overall increase in rates or changes in directions emerged. A constant assumption was adopted. International migration has been the most volatile component, affected by changes in legislation governing international migration. Geo-political changes in 2004 and 2009, when many Eastern European states joined the European Union (EU), resulted in increased immigration from the EU. From 2016, when the UK voted to leave the EU (the Brexit decision) immigration from the EU decreased and emigration increased. When the assumptions for the ETHPOP projection were set, it was uncertain what the future might bring. Therefore, three scenarios of high, medium and low net international migration between the UK and the EU were produced. This paper focuses on the medium scenario of modest reductions in immigration from the EU, compensated by modest increases in non-EU migration.

METHODS: Adjustment of projected ethnic populations to official population estimates
In our collaboration with the team at OHID, beginning in 2021, departures of components from component trends in 2011-2019 became apparent. In addition, the COVID-19 pandemic increased mortality and suppressed international migration

levels. The pandemic also affected fertility and internal migration moderately. The question the OHID asked the Leeds team was whether the projections could be adjusted to agree with official population estimates w at local and national scales. Such a revision required evaluation against the results of the end of decade census in 2021. ONS delivered a successful census in 2021: 97% of households submitted their forms (ONS 2021), of which 89% were online (ONS 2022).

One answer to the OHID question was to repeat the adjustment made for the 27 March 2011 Census base population to the 30 June 2011 mid-year population. The ETHPOP projected populations were adjusted to ONS mid-year populations for years between the two censuses. ONS publish population estimates by sex and age for local areas (LLTAs) for mid-years from 2011 to 2022. Adjustment of ETHPOP projected populations to ONS mid-year population estimates ensures agreement between the projected ethnic populations and the official estimates by age and sex. The adjustment was achieved using the following equation:

$$E_{esa}^{it}[2] = E_{esa}^{it}[1] \times \{\frac{M_{sa}^{it}}{\sum_e E_{esa}^{it}[1]}\} \qquad (1)$$

where E = ETHPOP projected population, M = MYEPOP estimated population, i = local area, t = mid-year, e = ethnic group, s =sex and a = age (five-year age group), [1] refers to the initial value of E and [2] to adjusted value of E. The adjustment ratios on the right-hand side are applied to each ethnic group. This procedure leaves the ethnic composition unchanged from the initial projected population.

### METHODS: Did use of the annual population survey (APS) help?
External members of the OHID review panel asked the research team to investigate addition of the Annual Population Survey as a constraint. The APS provides an independent estimate of the ethnic composition of the population for calendar years. However, the APS is seriously deficient: it only covers the household population, omits the population aged less than 16, produces credible numbers only at regional scale. For the calendar year 2011, APS population compositions by ethnicity fail to agree with the 2011 Census "gold standard". APS population estimates suggested, for example, that the Pakistani population, with high fertility rates and a large net immigration flow, may decrease across some years. Despite efforts to "guestimate" the missing information, we concluded that the APS could not fulfil a role as a third input.

FINDINGS: Evaluation of ETHPOP projected and adjusted populations against the 2021 Census for England
Full results to date of the ETHPOP projections can be found in Wohland et al. (2017) and partial results of the adjusted projected populations are available in the Open Science Framework repository (https://osf.io/wdqbj/). By June 2024 we will complete the adjusted estimates using re-based population estimates by ONS to be released in November. The evaluation here focuses on an England summary using roll-forward mid-year population estimates.

Table 1 compares ETHPOP projected populations against the 2021 Census populations adjusted to the MYEPOP estimates for mid-year 2021, created after the publication of the relevant census tables. The correlation between the two series of percentages is extremely close at 0.999 with a regression slope of 1.006. However, when we compare the percentage changes between 2011 and 2021 some differences emerge (Figure 1). The decrease in the White British and Irish group is larger between censuses than in the projection. This can be attributed to an over optimistic assumption about mortality decreases. In mid-decade the centuries long improvement in life expectancies ceased, flat-lining until 2019 and decreasing sharply in 2020 because of Covid-19 deaths and in 2021 because of continuing Covid-19 deaths and further excess deaths from other causes that were a consequence of delayed diagnosis and lengthening waits for hospital procedures. These factors affected all groups but were most acute for the majority group with a much older age structure than all minority groups bar the Black Caribbean, which had immigrated to the UK between 1948 and 1963. The decrease in the share of the White Other group reflects a downturn in immigration and an upturn in emigration since 2016 (the year of the Brexit referendum). Our assumption for international migration in the medium scenario incorporates this effect for immigration but our emigration assumption turned out to be a radical under-estimates. For the Chinese group the 2021 Census records a major decrease compared with the ETHPOP projection, due to students returning to China in 2020 because of the Covid-19 pandemic and missing the 2020-2021 academic year. In part, video communication, online teaching materials and emails substituted for international migration. The higher census share for the Other ethnic group reflects a diversification of the countries from which immigrants were recruited to the meet shortages in the UK labour market.

Figure 1: Comparison of the ethnic composition of the population of England at mid-year 2021: Estimates based on forecasts from 2011 and estimates based on the 2021 Census

ETHPOP and CENSUS Populations 2021
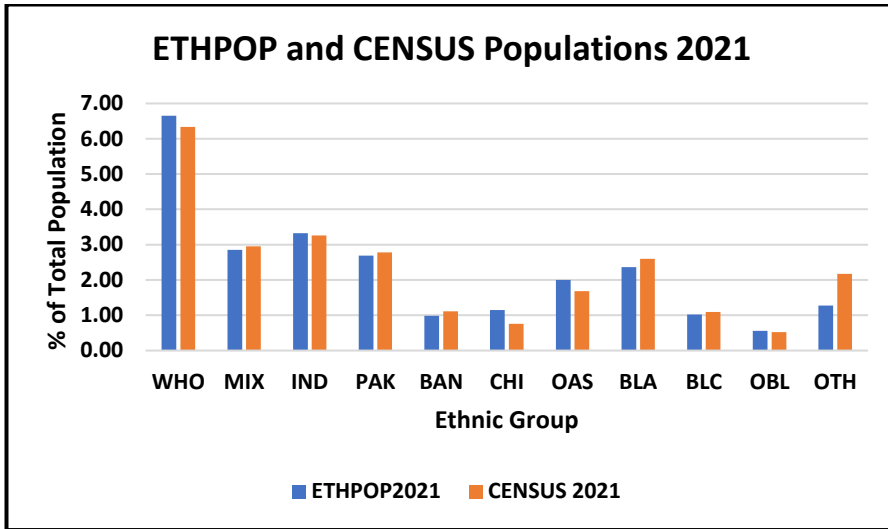
Notes: See Table 1 for the full names of ethnic groups
The percentages for the White British and Irish group are not plotted to show ethnic minority percentages clearly.
Sources: ETHPOP2021: Open Science Framework, https://osf.io/wdqbj/. Computed by the authors.
CENSUS 2021 ethnic populations: Office of National Statistics using the NOMIS portal to extract Table
RM032 - Ethnic group by sex by age - Nomis - Official Census and Labour Market Statistics (nomisweb.co.uk)

Two versions of the 2021 MYEPOP were produced by ONS, the first published in late 2022 before the 2021 Census results were available and the second in February 2023 after those results were published. Application of mid-year population constraint does not change the ethnic composition comparison of Figure 1, but it does change the overall population numbers and age structures of each ethnic group. To evaluate the impact of those constraints, we made use of an Age Structure Error (ASE) measure developed by Wilson (2022), adapted as follows:

$$ASE_e \; = \sum_{sa} |E_{esa} - M_{esa}| \, / \sum_{sa} M_{esa} \qquad\qquad (2).$$

We computed 2021 ASEs for each ethnic group and all groups combined for three cases: ETHPOP projected populations with no adjustment, with adjustment to the mid-year estimates for 2021 computed *before* release of the 2021 Census tables. Seven of the 12 groups have ASEs below 10%, two have ASEs between 10 and 20% and three have ASEs between 30 and 60% (Chinese, Other, Other Black and Other Asian). ASEs generally decrease as the constraint becomes closer to the 2021 Census result. These errors are quite large. Part of the blame can be assigned to differences between the assumptions and outcomes for the components of change and part to the effects of Covid-19 in 2020-21. However, a major factor in the errors may be the result from asking the ethnic question afresh in a new census in 2021. New populations, not present at the 2011 Census, will have made "virgin" choices. Surviving immigrants over the decade choose their ethnicity for the first time in a census. The parents of infants born in the 2011-2021 decade assign an ethnicity to their new offspring in the 2021 Census. In addition, older people within the majority ethnic group (WBI) will have died and younger people will have emigrated though again a majority of these will be White British and Irish. And people who survived the decade may have changed their minds and opted for a different ethnicity.

DISCUSSION: Results for local and health areas
In the first phase of the project, projected and adjusted population for LTLAs were converted to CCG health areas using a lookup table. With the publication of the 2021 Census results, we converted the 2019 LTLAs to 2021 geographies for all mid-years between 2011 and 2021. These estimates will be converted to Integrated ICSs. This work awaits publication of the re-based population estimates.

DISCUSSION: Using the methods from 2021 to 2031
To apply the methods discussed in this paper for a further decade, it will be necessary to carry out new projections based on the 2021 Census results. However, ONS have been working on the development of new and much more ambitious method, the Dynamic Population Model, using administrative data to construct a comprehensive statistical population register (ONS 2023). The populations estimated using this new set of methods will include classification by sex, age and ethnicity more reliably than prior roll forward cohort-component methods. No plans are yet announced for using DPM results as the base populations for projections, so there may be a role for a new ETHPOP projection re-based on the 2021 results and adjusted each year to agree with DPM results.

REFERENCES

Lomax, N., Norman, P., Rees, P. & Stillwell, J. (2013) Subnational migration in the United Kingdom: producing a consistent time series using a combination of available data and estimates. *Journal of Population Research*, 30, 265-288. DOI: 10.1007/s12546-013-9115-z

Lomax, N., Wohland, P., Rees, P., & Norman, P. (2019a) The impacts of international migration on the UK's ethnic populations. *Journal of Ethnic and Migration Studies.* DOI: 10.1080/1369183X.2019.1577726

Lomax, N., Wohland, P., Rees, P., & Norman, P. (2019b) Data and Methods for Estimating and Projecting International Migration by Ethnicity. SOM2: Supplemental Online Material for the paper Lomax et al. (2019a). DOI: 10.13140/RG.2.2.25275.05929. Available from: https://www.researchgate.net/publication/331554200_Data_and_Methods_for_Estimating_and_Projecting_International_Migration_by_Ethnicity_SOM_for_The_impacts_of_international_migration_on_the_UK%27s_ethnic_populations.

Nafilyan, V., Islam, N., Mathur, R., Ayoubkhani, D., Banerjee, A., Glickman, M., Humberstone, B., Diamond, I., & Khunti, K. (2021). Ethnic differences in COVID-19 mortality during the first two waves of the Coronavirus Pandemic: a nationwide cohort study of 29 million adults in England. *European Journal of Epidemiology*, 36, 605–617

Norman, P., Rees, P. & Wohland, P. (2014) The use of a new indirect method to estimate ethnic-group fertility rates for subnational projections for England. *Population Studies*, 68:1, 43-64, DOI: 10.1080/00324728.2013.810300 https://doi.org/10.1080/00324728.2013.810300

ONS (2012) Annual population survey (APS) QMI. Office for National Statistics. https://cy.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/annualpopulationsurveyapsqmi#:~:text=The%20Annual%20Population%20Survey%20%28APS%29%20is%20a%20continuous,important%20social%20and%20socio-economic%20variables%20at%20local%20levels.

ONS (2020) Coronavirus related deaths by ethnic group, England and Wales methodology. Coronavirus-related deaths by ethnic group, England and Wales methodology - Office for National Statistics (ons.gov.uk)

ONS (2021) 97 per cent of households respond to Census 2021. Press release, Office for National Statistics. https://census.gov.uk/news/97-per-cent-of-households-respond-to-census-2021

ONS (2023) Dynamic population model, improvements to data sources and methodology for local authorities, England and Wales: 2011 to 2022. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/methodologies/dynamicpopulationmodelimprovementstodatasourcesandmethodologyforlocalauthoritiesenglandandwales2011to2022

Rees, P., Wohland, P. & Norman, P. (2009) The estimation of mortality for ethnic groups within the United Kingdom. *Social Science and Medicine*, 69, 1592–1607, doi:10.1016/j.socscimed.2009.08.015.

Rees, P., Wohland, P., Norman, P. & Boden, P. (2011) A local analysis of ethnic group population trends and projections for the UK. *Journal of Population Research*, 28(2-3): 149-184. DOI 10.1007/s12546-011-9047-4

Rees, P., Wohland, P., Norman, P. & Boden, P (2012) Ethnic population projections for the UK, 2001-2051. *Journal of Population Research*, 29(1): 45-89. DOI: 10.1007/s12546-0111-9076-z.

Rees, P., Wohland, P., Norman, P., Lomax, N. & Clark, S. (2017) Population Projections by Ethnicity: Challenges and a Solution for the United Kingdom. Chapter 18, 383-408 in Swanson D (ed.) *Frontiers in Applied Demography*. Switzerland: Springer.

Rees P., Paul Norman P., Webb S., Dowden K., Jones A., Price G., Lugton C., Pailthorpe C. (2023) A commentary on ethnic population estimates for local authorities and NHS commissioning bodies in England, 2011 to 2021 V4. Report prepared for the Office of Health Improvement and Disparities, School of Geography, University of Leeds. https://osf.io/wdqbj/

Wilson, T. (2022). Preparing local area population forecasts using a bi-regional cohort component model without the need for local migration data. *Demographic Research* 46(32), 919-956. https://doi.org/10.4054/DemRes.2022.46.32

Wohland, P., Rees, P., Norman, P., Lomax, N. & Clark, S. (2017) NEWETHPOP – Ethnic population projections for UK local areas 2011-2061. [data collection]. http://doi.org/10.5255/UKDA-SN-852508