

# Communicating ethnicity data quality

Darren Stillwell<sup>1</sup>, Richard Laux<sup>2</sup>

<sup>1</sup>*Cabinet Office Equality Hub, London, UK*

<sup>2</sup>*Cabinet Office Equality Hub, London, UK*

## Abstract

The UK government is committed to levelling up opportunity and ensuring fairness for all. The Cabinet Office's Equality Hub is central to this commitment and one of Equality Hub's priorities is improving the quality of ethnicity data held by UK government departments. High quality data about outcomes for different ethnic groups is vital - it can inform effective government interventions to reduce unjustified disparities.

The Equality Hub effects improvements through developing and overseeing programmes of data quality work. We provide guidance to analysts in other departments collecting and analysing ethnicity data. This session describes the different ways we communicate this guidance. It focuses on three outputs.

First, our innovative Standards for Ethnicity Data. These standards can help improve the collection and reporting of ethnicity data, to improve understanding of ethnic disparities and increase the value of ethnicity statistics. They can lead to better decisions and outcomes for the public.

Second, our 'Methods and Quality Report' (MQR) series. These look at different aspects of the quality of ethnicity data. They provide guidance based on real world ethnicity data. MQRs are an important way that we alert users to different data quality issues. One MQR described disparities in a particular approach to policing - "stop and search" - between black and white people in England and Wales. If geographical differences are not taken into account, these disparities can be misleading. Another MQR compared ethnicity data provided by a third party with data provided by individuals. There is also an MQR on why detailed, granular ethnicity data is important.

Third, blog posts are another crucial way we communicate important data quality issues. We have discussed the importance of harmonised standards for ethnicity. One blog post outlined the difficulties of comparing international ethnicity data. There has been much internal and external user interest in our posts.

All these outputs package our own thinking on data quality issues for a wider audience. One of our guiding principles is transparency. It is the root of all data quality improvement. By being transparent and communicating our views about different aspects of data quality, we hope to encourage other departments to be clear on the strengths and limitations of their data as well.

Feedback we have received from internal and external users on our outputs has been positive. They have led to internal and external stakeholders holding us in high regard for our technical expertise.

**Keywords:** ethnicity, standards, equality, communication

## **1. Introduction**

### **1.1 The Cabinet Office Equality Hub**

The Cabinet Office Equality Hub (Equality Hub, 2024) was formed in early 2021. It is central to the UK government's commitment to levelling up opportunities, and ensuring fairness for all.

A priority for the Equality Hub is improving the quality of data about disparities. This includes data about the types of barriers different people face. The Equality Hub considers data for people with statutory protected characteristics (Legislation.gov.uk, 2010). It also considers data for socio-economic and geographic inequality.

The Equality Hub has a team of analysts in the Equality Data and Analysis Division (EDAD). EDAD supports the other areas of the Equality Hub by analysing disparities in outcomes between people with these different characteristics in the UK. For example, disparities between educational qualifications, or healthy life expectancy. EDAD also advises UK government departments on data quality issues.

One area EDAD advises departments on is the quality of ethnicity data.

### **1.2 Improving ethnicity data quality**

Having high quality data about outcomes for different ethnic groups is vital. It can inform effective government interventions to reduce unjustified disparities between different groups.

Ethnicity data has become important in recent years. A person's ethnicity is often collected in UK government datasets. The 2017 Race Disparity Audit (UK Government, 2017) identified many of these datasets. It showed disparities that affected different aspects of people's lives. This led to the creation of Ethnicity facts and figures (Race Disparity Unit, 2017). This website contains over 180 government datasets about people from different ethnic groups and has proven to be a valuable resource for users of ethnicity data.

Ethnicity data on Ethnicity facts and figures and elsewhere needs to be fit for purpose. This is so the government and other organisations can use the data to understand disparities and their causes.

Recent publications have reiterated the importance of ethnicity data quality, and how to improve it. These include the:

- Race Disparity Unit's Quality Improvement Plan (Race Disparity Unit, 2021)
- Quarterly reports on progress to address COVID-19 health inequalities (Race Disparity Unit, 2021)

- The Inclusive Britain report (UK Government, 2022)
- The Inclusive Data Taskforce report (UK Statistics Authority, 2021) and subsequent ONS implementation plan (ONS, 2021)

Communicating the quality of ethnicity data is vital. It means that users can understand the strengths and limitations of data. They can use data appropriately.

### **1.3 Methods of communicating data quality**

EDAD rarely collects its own data. Its role includes influencing UK government departments to improve the quality of their data. One way that EDAD does this is by communicating how to improve ethnicity data, and interpret and use it better.

This paper focuses on three ways that EDAD does this:

- the Standards for Ethnicity Data
- Methods and Quality Reports (MQRs)
- blog posts

They are different presentation formats that meet a wide range of user needs. We undertake wide engagement to understand the quality requirements of these users. This means we can highlight important data quality issues through our communications and increase and emphasise the utility of data and analysis.

One of the guiding principles is transparency. It is the root of all data quality improvement. By being transparent and communicating views about different aspects of data quality, EDAD encourages other departments to be clear on the strengths and limitations of their data as well. By doing this, as well as adding value to data and supporting users with their data quality issues, EDAD can also emphasise and increase trustworthiness in people who produce ethnicity data.

## **2. Standards for Ethnicity Data**

### **2.1 Background and publication**

The report of the Commission on Race and Ethnic Disparities (Commission on Race and Ethnic Disparities, 2021) recommended that the Equality Hub should work with the Office for National Statistics (ONS) and the Office for Statistics Regulation (OSR) to develop and publish a set of ethnicity data standards. Once published, the use of these standards would result in

improved understanding and information gathering, and reduce the opportunity for misunderstanding and misuse.

Inclusive Britain (UK Government, 2022) was the government's response to the report by the Commission on Race and Ethnic Disparities. To address the Commission's recommendation on data standards, Inclusive Britain action 6 said:

"To ensure more responsible and accurate reporting on race and ethnicity, the [Equality Hub] will, by the end of 2022, consult on new standards for government departments and other public bodies on how to record, understand and communicate ethnicity data."

EDAD consulted on a draft set of Standards in summer 2022. The consultation report noted there were 47 responses to the consultation (Race Disparity Unit, 2023). EDAD revised the Standards based on comments from respondents and published a final set of Standards (Equality Hub, 2023). This was alongside an Inclusive Britain update report.

## **2.2 Users of the Standards**

The Standards meet action 6 by describing good practice in ethnicity data:

- collection
- analysis
- reporting

The Standards apply to people in government departments or public bodies working with ethnicity data. We also encourage organisations commissioned by public bodies to use the Standards.

It is understandable that people have different approaches to collecting and analysing ethnicity data - such as the reasons for collecting the data, or the time and budget they have available. However, following the Standards as closely as possible can help more responsible use of ethnicity data by government departments. This will help improve the quality of data and the use of the data across government.

## **2.3 Alignment with the UK Code of Practice for Statistics**

The Code of Practice for Statistics outlines the principles that producers of UK official statistics should commit to; it differs from the European Statistics Code of Practice (Eurostat 2017) in its form, but its content is closely aligned. The Ethnicity Data Standards reflect the three pillars in the UK Code:

- trustworthiness - a product of the people, systems and processes within organisations that enable and support the production of statistics and data.
- quality - statistics fit their intended uses, use appropriate data and methods, and are not misleading.
- value - statistics and data are useful, easy to access, remain relevant, and support understanding of important issues.

This means the Ethnicity Data Standards align with the Code of Practice, themselves the definitive set of high-level standards for UK statistics. We encourage people to read the ethnicity Standards alongside the Code of Practice.

In addition, the Standards, and the MQRs and blogs described later are ways that we address indicators 6.4<sup>1</sup>, 7.1<sup>2</sup> and 7.2<sup>3</sup> of the European Statistics Code of Practice.

## **2.4. Content**

This section outlines some main considerations in the standards.

### **2.4.1 Harmonisation**

The Standards recommend using the Government Statistical Service harmonised classifications for ethnicity (Civil Service Analysis Function, 2011) to collect and report on data. This improves coherence and comparability with other datasets across government. We also recommend collecting ethnicity data in as much detail as possible, and publishing ethnicity data in as much detail as possible. This allows data producers to understand differences in outcomes for different groups. For example, in England and Wales there are significant differences in outcomes between people in the black African, black Caribbean, and the ‘other black’ groups in some topics. These differences can be masked when producing analysis for a black aggregate group.

### **2.4.2 Not using the phrase ‘BAME’**

In the UK, the acronym ‘BAME’ stands for the group of ‘black, Asian and minority ethnic’ people. The Standards recommend not using this phrase because it emphasises some ethnic

---

<sup>1</sup> “Information on data sources, methods and procedures used is publicly available.”

<sup>2</sup> “The overall methodological framework used for European Statistics follows European and other international standards, guidelines, and good practices, while constantly striving for innovation.”

<sup>3</sup> “Procedures are in place to ensure that standard concepts, definitions, classifications and other types of standards are consistently applied throughout the statistical authority.”

groups and excludes others. It also groups together all ethnic minority groups into one and it's not possible to understand differing outcomes for specific ethnic groups.

Similarly, we also provide advice on how to write about ethnic groups in the standards (UK Government, 2021).

### **2.4.3 Increasing representation of ethnic minority groups in data collections**

Ethnic minority groups can be underrepresented in surveys and other data collections. This might be due to response rates, or other factors. The Standards give some practical ways that representation might be increased. This includes sample boosts, incentives and good survey design.

### **2.4.4 Self-reporting ethnicity**

In general, self-reporting leads to better quality ethnicity data. The Standards recommend moving towards that as an aim. Sometimes though proxy (or "third party") reporting of ethnicity is necessary. For example, if there is a need for ethnicity data for young children.

### **2.4.5 Controlling for other factors in analysis**

The Standards recommend analysing ethnicity data by controlling for other demographic factors. This can help put figures into context. As an example, the ONS did some sophisticated work in this area during the COVID-19 pandemic to understand what other factors affected mortality rates for different ethnic groups (ONS, 2023).

Undertaking complex regression analysis is not always going to be possible. But putting differences between ethnic groups into context can be helpful for users. For example, differences might be due to where people in some ethnic groups live.

### **2.4.6 Reporting on missing ethnicity**

Understanding how much ethnicity data is missing in a survey or data collection is important. The Standards recommend that users report on the levels of missing ethnicity data. This includes reporting on the different classes of missing data like 'prefer not to say' or 'unknown'. Understanding the level of missing data can help users interpret it better.

## **2.5 Next steps**

We have been promoting the Standards across the government. We have presented to government departments on them and we are getting good feedback.

EDAD is also working with OSR to increase the application of the Standards. This could include making links to the Standards much clearer in the supporting material around the Code of Practice for Statistics. They might also form part of checks carried out by OSR that ensure different statistical series comply with the Code. Both such initiatives would reinforce the importance of complying with the Standards.

The final thing we will consider is whether to produce similar standards for other equality groups, such as disabled people.

### **3. Methods and Quality Reports (MQRs) and blog posts**

#### **3.1 MQRs**

The MQRs are an important way that we communicate different data quality issues. The MQRs cover:

- how to interpret and use data
- what we are doing to improve ethnicity data
- data quality issues for specific ethnic groups

The MQRs summarise the issue and provide recommendations for users. They are considered, definitive and often collaborative pieces. For example, for the MQR on 'Using relative likelihoods to compare ethnic disparities' (Race Disparity Unit, 2020a), EDAD worked with colleagues in the ONS, and the UK Government's Department for Health and Social Care and Department for Work and Pensions.

Another important MQR describes how the relative disparity between black people and white people in a particular aspect of policing - 'stop and search' - can be misleading if geographical differences are not taken into account (Race Disparity Unit, 2021). And we published an MQR about collecting and reporting on detailed ethnic groups - for example, the 19 groups covered in the 2021 Census for England and Wales (Race Disparity Unit, 2020b). This MQR summarises some of the quality issues with the use and presentation of data for aggregated and detailed ethnic groups. It sets out EDAD's position of using detailed ethnicity classifications for data analysis. This is necessary to understand differences between groups.

#### **3.2 Blog posts**

EDAD has published 15 blog posts on ethnicity data quality issues. Like the MQRs, the blog posts are a crucial way that EDAD impresses upon users the importance of different data

quality aspects. In particular, these blogs have promoted the use of harmonised ethnicity classifications through a dedicated series of blog posts:

- Making ethnicity data more coherent and comparable (Data in Government, 2019)
- Asking people about their ethnicity (Data in Government, 2021a)
- Why data harmonisation is important (Data in Government, 2021b)
- Harmonising decentralised data collections (Government Analysis Function, 2022a)

Other blog posts that have received much interest are one that outlined the difficulties of comparing international ethnicity data (Data in Government, 2022), and another that described the strengths and limitations of attributing ethnicity using a person's name (Government Analysis Function, 2022b).

Our blog posts are quick, dynamic and demonstrate progress in data quality improvements. They can also provide a sense of narrative, for example the series of posts on harmonisation described before.

EDAD has collected together MQRs and blog posts in our Ethnicity data quality resources page (Race Disparity Unit, 2022).

#### **4. Conclusion**

The outputs discussed here package EDAD's thinking on data quality issues for a wider audience. They do a different, but complementary job, where the whole is more than the sum of the parts. It has been a conscious effort by EDAD to increase our reputation for thought leadership and professional expertise. Ethnicity is a high profile and sensitive topic area, with very challenging data quality issues. It demands the very best in the way that EDAD communicates those issues. And by doing these things, and by EDAD being the best it can, we have gained a reputation for trustworthiness with customers and stakeholders.

Feedback received from internal and external users on the outputs has been positive. They have led to stakeholders holding EDAD in high regard for having high technical expertise.

And by communicating data quality issues better, users can use and interpret the data better. If ethnicity data are used and interpreted better, the UK government can develop better policies, and there can be better outcomes for citizens.

## References

- Equality Hub*. (2024, February 5). GOV.UK. <https://www.gov.uk/government/organisations/the-equality-hub>
- Legislation.gov.uk. (2010). *Equality Act 2010, Section 4*. <https://www.legislation.gov.uk/ukpga/2010/15/section/4>
- UK Government. (October 2017). *Race Disparity Audit*. <https://www.gov.uk/government/publications/race-disparity-audit>
- Race Disparity Unit (2017). *Ethnicity Facts and Figures*. <https://www.ethnicity-facts-figures.service.gov.uk/>
- Race Disparity Unit. (2021). *Quality improvement plan: government ethnicity data*. <https://www.gov.uk/government/publications/quality-improvement-plan-government-ethnicity-data/quality-improvement-plan-government-ethnicity-data>
- Race Disparity Unit. (n.d.). *Quarterly reports on progress to address COVID-19 health inequalities*. <https://www.gov.uk/government/collections/quarterly-reports-on-progress-to-address-covid-19-health-inequalities>
- UK Government. (2022, March). *Inclusive Britain: government response to the Commission on Race and Ethnic Disparities*. <https://assets.publishing.service.gov.uk/media/62321381e90e070edcb8d97c/Inclusive-Britain-government-response-to-the-Commission-on-Race-and-Ethnic-Disparities.pdf>
- UK Statistics Authority. (2021, September 21). *Inclusive Data Taskforce*. <https://uksa.statisticsauthority.gov.uk/publication/inclusive-data-taskforce-recommendations-report-leaving-no-one-behind-how-can-we-be-more-inclusive-in-our-data/>
- Office for National Statistics. (2021). *Inclusive Data Taskforce Implementation Plan*. [https://uksa.statisticsauthority.gov.uk/wp-content/uploads/2022/01/1697-Inclusive-Data-Taskforce-IDTF-Implementation-plan-English-Web\\_v1-00.pdf](https://uksa.statisticsauthority.gov.uk/wp-content/uploads/2022/01/1697-Inclusive-Data-Taskforce-IDTF-Implementation-plan-English-Web_v1-00.pdf)
- Commission on Race and Ethnic Disparities. (2021, March 31). *The report of the Commission on Race and Ethnic Disparities*. <https://www.gov.uk/government/publications/the-report-of-the-commission-on-race-and-ethnic-disparities>
- Race Disparity Unit. (2023, April 17). *Government response to the consultation on standards for ethnicity data*. <https://www.gov.uk/government/consultations/standards-for-ethnicity-data/outcome/government-response-to-the-consultation-on-standards-for-ethnicity-data>
- Equality Hub (2023, April 17). *Standards for ethnicity data*. <https://www.gov.uk/government/publications/standards-for-ethnicity-data/standards-for-ethnicity-data>
- UK Government (2023, April 17). *Inclusive Britain: Update report*. <https://www.gov.uk/government/publications/inclusive-britain-update-report/inclusive-britain-update-report>
- UK Statistics Authority. (Updated 2022, May 5). *Code of Practice for Statistics*. <https://code.statisticsauthority.gov.uk/>
- Eurostat. (2017). *European Statistics Code of Practice — revised edition 2017*. <https://ec.europa.eu/eurostat/web/products-catalogues/-/ks-02-18-142>
- Civil Service Analysis Function. (2011, August 1). *Ethnicity harmonised standard*. <https://analysisfunction.civilservice.gov.uk/policy-store/ethnicity-harmonised-standard/>
- UK Government. (2021, December). *Writing about ethnicity*. <https://www.ethnicity-facts-figures.service.gov.uk/style-guide/writing-about-ethnicity/>
- Office for National Statistics. (2023, February 22). *Updating ethnicity contrasts in deaths involving the coronavirus (COVID-19), England and Wales*.

<https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/updatingethniccontrastsindeathsinvolvingthecoronaviruscovid19englandandwales/24january2020to23november2022>

Race Disparity Unit (2020a, August 18). *Using relative likelihoods to compare ethnic disparities.*

<https://www.gov.uk/government/publications/using-relative-likelihoods-to-compare-ethnic-disparities/using-relative-likelihoods-to-compare-ethnic-disparities>

Race Disparity Unit (2021, March 31). *Stop and search data and the effect of geographical differences.*

<https://www.gov.uk/government/publications/stop-and-search-data-and-the-effect-of-geographical-differences/stop-and-search-interpreting-and-describing-statistics>

Race Disparity Unit (2020b, December 22). *Ethnicity data: how similar or different are aggregated ethnic groups?*

<https://www.gov.uk/government/publications/differences-in-the-quality-of-ethnicity-data-reported-by-individuals-and-third-parties>

Data in Government. (2019, June 4). *Making ethnicity data more coherent and comparable.*

<https://dataingovernment.blog.gov.uk/2019/06/04/making-ethnicity-data-more-coherent-and-comparable/>

Data in Government. (2021a, June 15). *Asking people about their ethnicity.*

<https://dataingovernment.blog.gov.uk/2021/06/15/asking-people-about-their-ethnicity/>

Data in Government. (2021b, January 29). *Why data harmonisation is important.*

<https://dataingovernment.blog.gov.uk/2021/01/29/why-data-harmonisation-is-important/>(<https://dataingovernment.blog.gov.uk/2021/01/29/why-data-harmonisation-is-important/>)

Government Analysis Function. (2022a, April 20). *Harmonising decentralised data collections.*

<https://analysisfunction.civilservice.gov.uk/blog/harmonising-decentralised-data-collections/>

Data in Government. (2022, January 25). *Comparing ethnicity data for different countries.*

<https://dataingovernment.blog.gov.uk/2022/01/25/comparing-ethnicity-data-for-different-countries/>

Government Analysis Function. (2022b, April 20). *Attributing ethnicity using someone's name.*

<https://analysisfunction.civilservice.gov.uk/blog/attribution-ethnicity-using-someones-name/>

Race Disparity Unit. (2022, January 4). *Ethnicity data: data quality resources.*

<https://www.gov.uk/government/collections/ethnicity-data-methods-and-quality-reports>