Understanding Patient Data

UNIVERSITY of WORCESTER

ETHNIC CATEGORIES AND CATEGORISATION OF GYPSY, ROMA, AND TRAVELLER COMMUNITIES

Desk review process and key themes

January 2025

Prepared By

Dr Peter Unwin, Josie O'Driscoll, Allison Hulmes, Stacey Hodgkins, Shamus McPhee, Dr Rosa Cisneros, Jackie Bolton and Prof. Győző Molnár

Contents

Introduction and background

A brief historical background to Gypsy, Roma, and Travellers

Nomadic people were probably well established in Britain by about 1000 A.D. and these included English Gypsies, Welsh Gypsies, Scottish Gypsy Travellers, and Irish Travellers. In the late 1400s Roma people began arriving in Britain, having originated in India (Travellers Times, 2019). The categorisation of groups is complex and contested, some groupings preferring the word 'Romany' to 'Gypsy' as in 'English Romany' and 'Welsh Romany', and sometimes the word 'Roma' is used to include all the above cultures under the one umbrella term. 'Romani' is a term also sometimes used to represent English Gypsies, Welsh Gypsies, Scottish Gypsy Travellers, and Irish Travellers. 'Gypsy' is regarded as a derogatory word in much of Europe but in Britain it is an acceptable term (McLaughlin, 2008). The abbreviation 'GRT' is not seen by communities as an appropriate term as it homogenises their distinct cultures, instead of recognising their differences (Hulmes & Unwin, 2024). Over the centuries all these groups have suffered various forms of discrimination as their fortunes and lifestyles have changed. Only a very small minority are now nomadic, partly because modern workforces, such as in agriculture, no longer require large numbers of itinerant workers and legislation, such as The Police Criminal, Sentencing and Courts Act (2022) has sought to criminalise the nomadic way of life (Hulmes & Unwin, 2024).

Roma tend to be 'settled' once in Britain and live in 'bricks and mortar', usually concentrated in cities and large towns with greater employment opportunities. The ONS Census (2021) reported that, in England and Wales, 78% of English Gypsies, Welsh Gypsies, Scottish Gypsies and Irish Travellers also now live in 'bricks and mortar' with 22% living on sites in chalets or caravans. There are also a very small number of Romani who live 'roadside' and do still travel year-round.

Rationale for the review

The need for accurate data collection has been noted by many reports in relation to the health of Gypsies, Roma, and Travellers (e.g., House of Commons Women & Equalities Committee, 2019; Unwin et al, 2022; GATE Herts, 2024; UK Health Data Research Alliance, 2024), yet only limited advancement has been made in this area. Whilst the 2021 National Census included 'Roma' as distinct ethnic category for the first time, alongside a combined 'Gypsy or Traveller' option, NHS data collection relating to patients' cultural identity has remained dated despite recommendations by a report commissioned by Inclusion Health (Aspinall, 2014).

To gain an accurate sense of the health needs of Gypsy, Roma, and Traveller communities at both national and local levels has been challenging because of their absence in standardised datasets. Obtaining such information is seen as crucial if unmet health needs are to be successfully addressed. Effective service development and deployment is seen as ineffective without better quality and more granular information on the range of the specific health needs of Gypsies, Roma, and Travellers, whose lifestyles are often very different to most of the UK's population. Aspinall (2014) also noted the significant inconsistencies in the cultural classifications being used across a range of NHS services in England and across the UK, which remain in existence. For instance, the General Practitioner Extraction Service (GPES) data set has 18 ethnic categories (based on the 2011 Census) whereas the Hospital Episode Statistics (HES) data set only contains 16 ethnic categories. HES contains no Gypsy, Roma, or Traveller categories and GPES has only a Traveller category. Consequently, data on Roma is absent in both GPES and HES, and data on Gypsy and Travellers are absent in HES and minimal in GPES (ONS, 2024).

Given the lack of health data around Gypsy, Roma, and Traveller communities and both dated and ineffective ethnic categorisations across health services in the UK, we have conducted a scoping review that focused on identifying existing relevant sources that centred on ethnic categorisation and associated factors. Scoping reviews enable researchers to examine central issues surrounding a research area and discover key sources and types of evidence available (Tricco et al., 2018), without being restricted by a potentially narrow range of quality-defined studies and by different methodological approaches used. In accordance with scoping review recommendations (Peters et al., 2015), we used broad research questions combined with clear definitions of the concepts relevant to the study's scope. We have established three main questions to inform the review, which are as follows:

1. What research has been conducted that has examined the importance of ethnic categorisation in demographic data?

2.What are the main approaches to determining ethnic categories?

3. What research has specifically focused on the ethnic categorisations of Gypsy, Roma, and Traveller populations?

Method

We selected and searched the following databases: ProQuest Central, CINAHL, MEDLINE and Web of Science. For ProQuest, we selected four relevant databases: Public Health, Sociology, Political Science and Social Science to allow for the capture of a broad range of disciplinary perspectives on ethnic categorisation. To identify relevant sources, we developed and deployed the following search term syntax for use across all databases identified: ethnic category, ethnic classification, ethnic categorisation, ethnic monitoring, ethnicity data, ethnic categorisation of Gypsy, Roma, and Travellers. From the search syntax, the following search word sequences was developed and used to search the title or abstract:

(title ("ethnic classification" OR "ethnic group terminology" OR "ethnic categories" OR "ethnicity classification") OR abstract ("ethnic classification" OR "ethnic group terminology" OR "ethnic categories" OR "ethnicity classification") AND (gypsy OR gipsy OR roma OR traveller OR "gypsy roma or traveller")) AND PEER (yes)

The search was restricted to peer-reviewed sources in English with full text availability. The literature search's cut off point was 30th May 2024 and considered for inclusion anything that was published up until that date. Overall, 342 sources were identified in four databases with 16 duplicates removed, leaving 328 sources for initial analysis (for breakdown of sources by database please see Appendix 1). Out of 328 sources, 27 were retained for further analysis with the view to be included in the scoping review. After an in-depth review of all 27 sources, six were eliminated and 21 included in the final review. Appendix 1 depicts the process of the review in the form of a flow chart. The reference list of all 27 sources that were selected for full analysis can be found in Appendix 2.

In addition to sources identified through search terms and databases, we also asked the members of the project Advisory Group (AG) to identify and suggest relevant sources that should be considered for inclusion in the scoping review. AG members identified 14 additional sources (Appendix 3), out of which 13 were included in the review.

All reviewed sources (41 in total, out of which six were eliminated during the detailed review process) have been included in tables. We separated sources identified by database search and by members of the AG. A summative content of all sources included in the scoping review can be found in **Tables 1 and 2** in the additional materials. The review has identified the following key themes which will be elaborated on below:

- Challenges of defining race and ethnicity;
- Rationale for collecting ethnic data;
- Modes of recording ethnic identity;
- Key issues with ethnic categories in UK health data.

Themes/findings

Challenges of defining race and ethnicity

The terms ‘race’ and ‘ethnicity’ are often used interchangeably, although they are argued to refer to different aspects of humankind. Traditionally, race tends to refer to biological (often visible) characteristics, based on which different groups of humans have been, and can be, differentiated. It is, however, essential to recognise that race and ethnicity are not inherently linked to biology or genetics. Instead, race and ethnicity are social constructs shaped by historical and cultural factors (Lewis et al., 2023). Many social scientists use ‘race’ to express that the concept is socially constructed as one’s race influences one’s life experiences. Ethnicity, alternatively, is argued to signal cultural belongingness expressed though, for instance, language, traditions, and shared ancestry. Salo (1979) explains that ethnic identity refers to the categorisation of individuals as members of a particular ethnic group based on various criteria. However, there are no universal standards for defining the cultural aspects of ethnicity, and they tend to change over time and in cultural contexts.

Lewis et al. (2023) note that the terminology regarding descriptions of race and ethnicity has not been consistent and has evolved over the years. Differences in terminology, methods of data collection, individual perceptions of group identity, and changing demographics present challenges in determining racial and ethnic categories that are specific and acceptable to all individuals. Lewis et al. (2023) also observe that there is significant variation in how individuals and communities self-identify and how healthcare providers and researchers define and report these categories. Therefore, standardisation of terminology is essential for promoting clarity, accuracy, and inclusivity in medical literature and reporting while keeping in mind sensitivities and controversies related to race and ethnicity.

Data on race and ethnicity tend to be more reliable when self-reported (Lachowsky et al., 2020). However, it is pertinent to note that even self-reported identities can change over time depending on one’s affiliation with their community, perceptions of acceptance, advantages or disadvantages, and safety concerns. In essence, one’s ethnicity is not an ingrained or permanent trait, instead it is influenced by social and environmental factors over time (Lachowsky et al., 2020).

Rationales for collecting ethnic data

Reasons for collecting data on ethnicity have been associated with benefits for society such as encouraging equality, reducing discrimination, enhancing policy effectiveness, and increasing diversity. Conversely, Sookrajowa (2021) noted that by abridging, standardising, and surveying societies, states have tried to make people more observable and, thus, easier to express power over them. The core of Sookrajowa’s observations can be viewed as an adaptation of Foucault’s (1979) concept of Panopticon. Foucault held that governments in contemporary societies exert their power over people through various forms of surveillance. Such surveillance might take place through the actual monitoring of people through CCTV cameras or various forms of bureaucracy. As people are under constant surveillance, they internalise the social gaze of authority which leads them to regulate their own actions. In essence, observing and keeping data on people may be perceived as a form of social control.

Based on the work of Rallu, Piché, & Simon (2004, see also Simon, Piché & Gagnon, 2015), Morning (2008) and Sookrajowa (2021) suggest a four-fold typology for collecting data on a population’s ethnic categories. These are: a) counting to control, b) not counting to support

national integration, c) counting or not counting in the name of national hybridity and d) counting to justify positive action.

a) *Counting to control* – this reason relates to colonial times and stems from the need to effectively exert and retain power over indigenous populations. This practice, particularly prominent in the 19th and 20th centuries, shows how data collection served as a tool for domination when used for hierarchical classification, social division, and resource control at the expense of indigenous peoples.

b) *Not counting to support social integration* – this approach follows the logic that collecting and monitoring people's ethnic background has the potential to trigger division and separation between different communities. By not monitoring ethnicity, people's focus shifts from differences to commonalities, which, in turn, encourages a cohesive cultural identity that transcends ethnic lines. The argument for not collecting ethnic data is perceived to be associated with reducing the risk of stereotyping, essentialising and ethnic profiling.

c) *Counting or not counting for national hybridity* – in multicultural societies, the need often arises to tailor services according to the ethnic makeup of the society. In such societies, ethnic data is argued to help reveal social disparities and emerging hybrid identities. Through ethnic monitoring, we can better understand where inequalities exist which aids policymakers to create specific initiatives to support underrepresented or marginalised communities to ensure equitable access to resources and opportunities. Ethnic data can also be used to help cement ideologies associated with multiculturalism and celebrate the cultural multifacetedness that is present across the social spectrum in a quest for common goals and shared values which are believed to strengthen national cohesion. This approach emphasises inclusivity and reflects society's growing diversity and intersectionality. There is also a belief that not collecting multifaceted ethnic data that would reflect hybrid identities helps with national belongingness, as this encourages individuals to see themselves as members of a unified nation, rather than as members of separate ethnic groups. This latter approach follows a simplified form of data collection by assigning individuals to a single ethnic category, often based on self-identification or prioritizing one aspect of their heritage.

d) *Counting to justify positive action* – The main aim of this type of ethnicity data collection is to combat racism and related discrimination across the social spectrum. Positive action-centred ethnicity data collection derives from the idea that statistical data on ethnicity will help inform social and governmental actions to address existing discrimination and inequalities thereby improving the overall quality of life across society. When inequalities and marginalisation are identified then relevant bodies can spring into action to develop policies and social initiatives designed to curb existing ethnicity-based discrimination by addressing institutional and other forms of racism. In line with the idea of positive action, policies are created that prevent active discrimination and launch initiatives that promote integration with the view to reduce (and ultimately close) the gap between different ethnic groups in society by increasing access to resources and opportunities.

Given the historical atrocities experienced by members of the Gypsy, Roma, and Traveller communities (e.g., Hulmes & Unwin, 2024), whereby their predecessors' and their own ethnic identity was collected for the purposes of hierarchical classification, it may explain some of the issues behind the limited ethnicity data on those communities. For instance, Gypsy, Roma, and Traveller communities may be more likely to perceive data collection as *counting to control* because of the discrimination and persecution they experienced in other contexts due to having their ethnicity identified. For these reasons, many Gypsies, Roma, and Travellers prefer a *not counting to support social integration* approach to ethnicity data collection. In fact, members of those communities have expressed that declaring their

ethnicity in healthcare settings felt like a precursor to receiving a different, suboptimal treatment or care (See **full research report).**

Modes of recording ethnic identity

There are multiple modes of recording ethnic data such as observational, community-based and self-identification approaches (Ringelheim, 2011). Ahmed, Feliciano & Emigh (2007) note that external classification can conflict with self-identifications. In other words, external labelling may miss subtle cultural differences that outsiders do not recognise, but insiders wish to preserve; or outsiders can highlight differences that insiders wish to eradicate thereby creating/exaggerating ingroup disparities. For instance, in a healthcare setting, external classification may take place when a patient's ethnic identity is entered by a staff member based on their observation and the ethnicity category entered is not confirmed by the patient. That is, an individual born and raised in the UK may identify as Gypsy due to their parents' cultural heritage. When the individual chooses "Gypsy" on a survey, their self-identification accurately reflects their identity. However, if data was collected by external observer, they might incorrectly categorise the individual as White British, leading to inaccurate data representation. Therefore, Ringelheim (2011) notes that self-identification is the most in-line with respect for individuals' rights (see also Benett, 1997; Felouzis, 2010) and should be the practice across all healthcare settings.

Nevertheless, self-identification may be an issue when data are collected for anti-discrimination purposes. In such instances, data is collected to identify, address and prevent discrimination based on a protected characteristic, e.g., ethnicity, age, gender, race, etc. However, an individual may not always identify in the way they are perceived by others, or they strategically misreport their ethnic identity to avoid stigma. For instance, Gypsies and Travellers often do not declare their true ethnic identity to avoid discrimination in the health services (See **full research report**). This can bias data collection and the number of people using health services from those communities. Under such circumstances it may be justified to deploy other modes of ethnic identification such as external classification (e.g., observational). However, as Ringelheim (2011) suggests, when other modes of ethnic identification are used, it should be ensured that affected individuals have input into the categorisation process. Self-identification of ethnicity can pose further challenges, which may include, for instance, changes to identities over time, individual interpretations of ethnic categories, the acquisition of multiple ethnic identities, and the hiding of identity due to fear of stigma and discrimination.

Despite the challenges, sources included in the review and the latest White Paper on *Enhancing diversity and quality in health data* (Quattroni et al. 2024) recommended using self-reporting for ethnic data collection whenever possible.

Key issues with ethnic categories in UK health data

Some of the sources directly focus on key issues with ethnic categories in the UK health data landscape. Specifically, Office for National Statistics (ONS) reports (2022, 2024) identify concerns which will require modifications to the current system. For instance, whilst the 2021 Census included 19 ethnic categories (see Appendix 4, list 1), ethnic categories in the NHS in England are out of date. The General Practitioner Extraction Service (GPES, see Appendix 4, list 2) data set has 18 ethnic categories (based on the 2011 Census) whereas the Hospital Episode Statistics (HES, see Appendix 4, list 3) data set only contains 16 ethnic categories. HES contains no Gypsy, Roma, or Traveller categories and GPES has only a

Traveller category. Consequently, data on Roma is absent in both GPES and HES, and data on Gypsy and Travellers are absent in HES and minimal in GPES (ONS, 2024).

Public Health Scotland deploys an 18-item list of ethnic categories (see Appendix 4, list 4) that includes Gypsy/Traveller (combined category), Roma, Showman/Showwoman (combined category). While the 18-item list used in Scotland is the most advanced in the UK, it still combines Gypsy and Travellers as one category which is not in line with recommendations by charities such as the Traveller Movement (No date). Gypsies and Travellers are identified within the four groups who are the most discriminated in Scotland (Scottish Government, 2015). Discrimination is the reality for many Gypsies and Travellers, despite the Scottish Government (2015, p. 4) having reported that nearly 7 in 10 (69%) Scots felt that 'Scotland should do everything it can to get rid of all kinds of prejudice'.

Public Health Wales (see Appendix 4, list 6) deploys an ethnic category list containing 16 items. It has a combined category of Gypsy or Irish Traveller, but the current list is not in-line with the recommendations made by Welsh Government (2018) which recognise that Romani Gypsies, Irish Travellers, and migrant Roma are ethnic groups under the Equality Act 2010. The Welsh Government's (2022) anti-racist strategy clearly recognises fundamental social issues with how Gypsy and Traveller communities are treated. However, in their 2022 report, they still refer to Gypsies and Travellers as one ethnic category, which is contrary to recommendations by the Traveller Movement (No date) and their own 2018 recommendations (Welsh Government, 2018), which state: 'Romani Gypsies, Irish Travellers and migrant Roma are ethnic groups under the Equality Act 2010.'

Ethnic categories in Health and Social Care in Northern Ireland (see Appendix 4, list 6) contain 11 items, which makes it the shortest one in the UK, but it does contain Irish Traveller as an individual category. It, however, appears that ethnic categories used in health care do not reflect the ethnic composition of the population of Northern Ireland. For instance, based on the 2021 Northern Irish Census (NISRA, 2021), the Filipino population is larger than their Irish Traveller population and they do not have their ethnicity listed. Also, whilst there is no significant difference between Roma and Pakistani populations, the Roma ethnic category is absent from the health care list of ethnic categories.

Main recommendations

a) **Need for Contextual Data Collection** - Ethnic identity is often connected to other identities such as education and socio-economic status (see Kisfaulidy, 2018) and, thus, it should be collected along with other key demographic variables to be meaningful. Contextual data has the potential to enhance our understanding of population dynamics by providing insight into the factors that shape disparities, trends, and outcomes. For instance, identifying the socio-economic status and education background of Gypsies, Roma, and Travellers may aid with addressing exiting structural inequalities.

b) **Recognition of Fluid Nature of Identity** - Ethnic identity is not stable or monolithic. It needs to be recognised that individuals' identity may shift between or span across categories over time (e.g., identity pluralisation). In other words, people's identities can change due to life experiences, exposure to new cultures, personal growth, and shifts in societal attitudes. Individuals may identify differently in their youth compared to later in life, based on their evolving understanding of culture, heritage, or personal experiences. For instance, the importance of cultural belonging may increase for Gypsies, Roma, and Travellers as they age and improve their knowledge of and appreciation for their own ethnic background.

c) **Need for Adaptive Ethnic Categories** - As populations change, cultures develop, and new social dynamics emerge, the way ethnicity is understood and categorised should reflect these transformations. Adaptive ethnic categories allow for more accurate, inclusive, and context-sensitive data collection and better understanding of human diversity. Given that ethnic identity is shifting, and changeable, ethnic categories should reflect social changes and be revisited (and revised if deemed necessary) in regular intervals. For instance, over time identities transform by both gaining new and losing old elements, which might explain why Gypsy Travellers in Scotland incorporated Scottishness into their identity and prefer to be called Scottish Gypsy Travellers.

d) **Increased Specificity in Health Data** - Pertaining to the focus of this review, the following ethnic categories should be separated when health-related ethnic data are collected: Gypsy, Roma, and Traveller[1] (Please see the **Gypsy, Roma & Traveller communities: equitable data collection report** where we present six empirical evidence-based categories to be used across health services in the UK to replace the current classifications);

e) **Need for Expanded Ethnic Categories** - In the UK health services, a minimum of 19 ethnic categories should be deployed across the sector, in line with the latest recommendation on ethnic categories in the health service by Quattroni et al. (2024). Whilst this source was not part of the review, given the dated ethnic categories used across the UK health services, the recommendation is pertinent to note.

f) **Positive Purpose Driven Data Collection** – Whilst ethnicity data can be and has been collected for multiple reasons and Morning (2008) and Sookrajowa (2021) have detailed a four-fold typology for ethnicity data collection, it is recommended that ethnicity data should only be collected to promote positive social action and that the rationale for data collection to be clearly and effectively communicated to the population.

g) **Need for Standardised Terminology** - Standardisation of terminology of ethnic categories is essential. By having consistent and generally understood terms, we can ensure

[1] Please note that in Scotland the term Gypsy Traveller has been in used since 2008, in which year the Scottish Government recognised members of the community as an ethnic group. Therefore, we would recommend that in Scotland the term Scottish Gypsy Traveller is used as a single ethnic category.

that the information gathered is accurate, comparable, and useful in addressing social disparities for promoting clarity, accuracy, and inclusivity in medical literature (Lewis et al., 2023). For instance, standardised terminology would allow for consistency in how ethnic categories are defined and used across different surveys, studies, and datasets. This ensures that data can be accurately compared across time, locations, and research projects.

j) **Self-Reporting Preferred** - Self-reporting for ethnic data collection should be used as a form of data collection whenever possible. Allowing individuals to define their own ethnic identity provides more reliable and meaningful data, while also ensuring that individuals are empowered to express themselves as they see fit.

Conclusion

Ethnic identity is fluid, multifaceted and often connected to socio-cultural factors such as education and socio-economic status, necessitating its collection alongside other demographic variables. As ethnic identity is dynamic and can shift over time, categories should be regularly checked and, if necessary, updated to reflect socio-cultural changes. For instance, the frequently used 'GRT' acronym to represent people from Gypsy, Roma, and Traveller communities is now considered inadequate and essentialising. For better health provisions and outcomes for Gypsy, Roma, and Traveller communities, more granular ethnic data are essential. Therefore, in the UK, a minimum of 19 ethnic categories should be used across health services to capture the ethnic diversity of the population. Furthermore, it needs to be ensured that ethnicity data is self-reported and exclusively collected to support positive social action, with standardised terminology ensuring clarity and inclusivity.

This scoping review has highlighted some key typologies, theories, and recommendations around ethnicity data collection in general and as those relate to Gypsies, Roma, and Travellers, in particular. The review has also supported, complemented and informed the empirical work we carried out on behalf of Understanding Patient Data with Gypsy, Roma, and Traveller communities to uncover and unpack their perspectives and experiences in relation to ethnicity data in health care. Please see the **full research report here**.

References

Aspinall, P. (2014). *Hidden Needs Identifying Key Vulnerable Groups in Data Collections: Vulnerable Migrants, Gypsies and Travellers, Homeless People, and Sex Workers*. London. Inclusion Health.

Foucault. M. (1979). *Discipline and Punish: The Birth of the Prison*. London: Penguin Books.

GATE Herts (2024). *Data Inclusion Briefing*. Available at: https://gateherts.org.uk/wp-content/uploads/2024/04/DATA-Grouping-Briefing-2-1.pdf

House of Commons Women and Equalities Committee (2019). *Tackling inequalities faced by Gypsy, Roma, and Traveller communities Seventh Report of Session 2017–19.* Available at: https://publications.parliament.uk/pa/cm201719/cmselect/cmwomeq/360/360.pdf

Hulmes, A. & Unwin, P. (2024). Rural Social Work with Romani and Travellers. In: Pye, J. & Turbett, C. (eds), *Rural Social Work in the UK and Beyond: Themes and Challenges for the Future*. London: Palgrave Macmillan.

NISRA (2021). *Main statistics for Northern Ireland Statistical bulletin: Ethnic group*. Available at: www.nisra.gov.uk/system/files/statistics/census-2021-main-statistics-for-northern-ireland-phase-1-statistical-bulletin-ethnic-group.pdf

ONS (2024). *Quality of ethnicity data in health-related administrative data sources by sociodemographic characteristics*, England. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthinequalities/articles/understandingconsistencyofethnicitydatarecordedinhealthrelatedadministrativedatasetsinengland2011to2021/may2024

Page, M.J., McKenzie, J.E, Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D. et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372:71.

Peters, M.D.J., Godfrey, C.M., Khalil, H., McInerney, P., Parker, D., & Soares, C.B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, 13(3), 141–146.

Quattroni et al. (2024). Enhancing diversity and quality in health data. *Zenodo*. https://doi.org/10.5281/zenodo.13890530

Rallu, J.-L., Piché, V., & Simon, P. (2004). De´mographie et Ethnicite´: Une Relation Ambigue¨. In G. Caselli, J. Vallin, & G. Wunsch (Eds.), *De´mographie: Analyse et Synthe`se* (pp. 481–515). Paris, France: Institut National d'Etudes De´mographiques.

Simon, P., Piché, V., Gagnon, A.A. (2015). The Making of Racial and Ethnic Categories: Official Statistics Reconsidered. In: Simon, P., Piché, V., Gagnon, A. (Eds.), *Social Statistics and Ethnic Diversity*. IMISCOE Research Series. Springer, Cham. https://doi.org/10.1007/978-3-319-20095-8_1

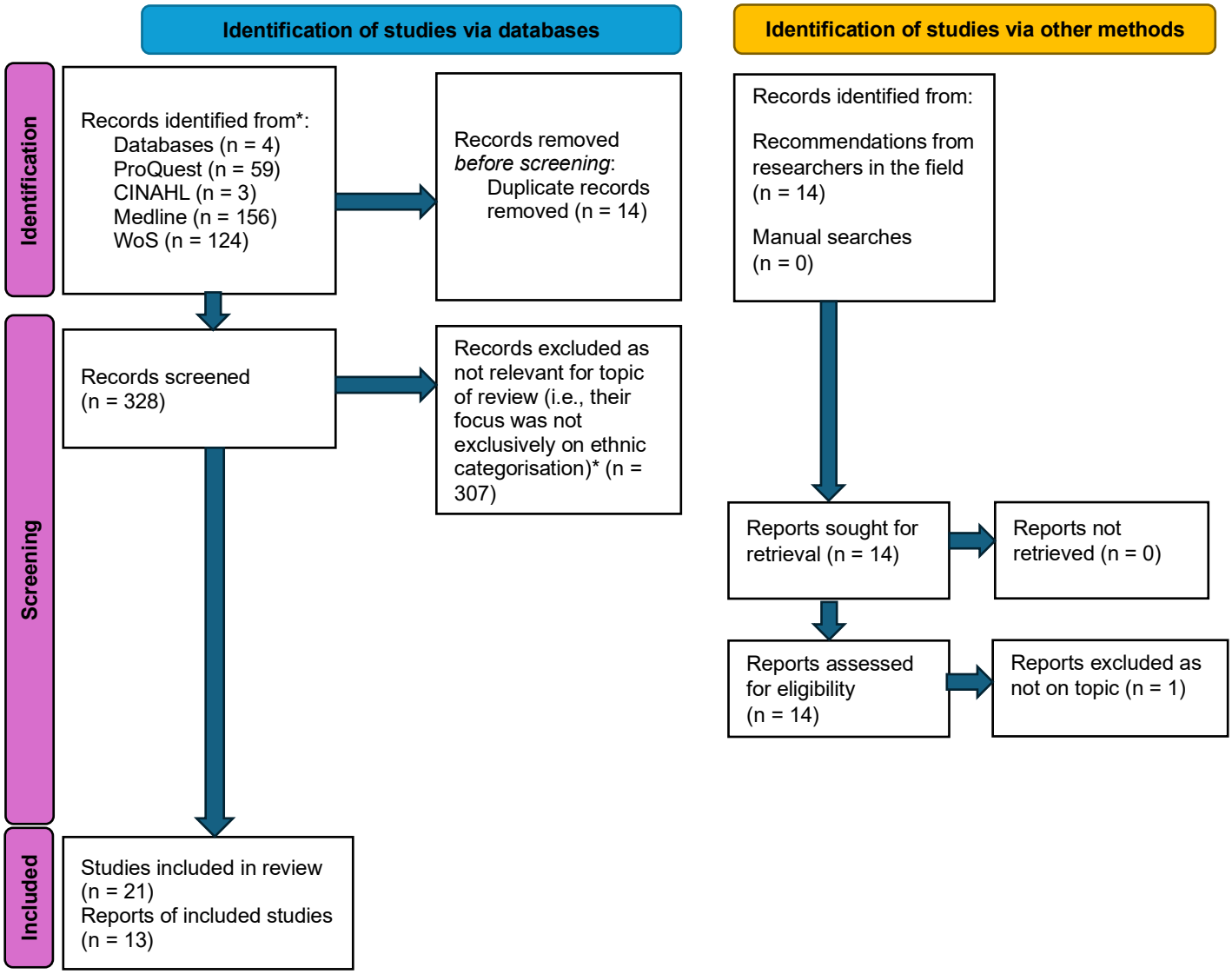
Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., Lewin, S., Godfrey, C.M., Macdonald, M.T., Langlois, E.V., Soares-Weiser, K., Moriarty, J., Clifford, T., Tuncalp, O. & Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Annals of internal medicine, 169(7), 467–473. https://doi.org/10.7326/M18-0850

UK Health Data Research Alliance (2024). *Enhancing diversity and quality in health data*. Available at: https://zenodo.org/records/13890530

Unwin, P. & Price, L. (2022). 'Being a Gypsy'. in Hughes, M.(ed.), *Social Exclusion in the UK*. London: Critical Publishing Ltd.

Appendices

Appendix 1. Scoping review flow chart



*All records were excluded by a human and no automation tools were used for selection.

Appendix 2. Sources identified via databases

1) Ahmed, P., Feliciano, C., & Emigh, R.J. (2007). Internal and External Ethnic Assessments in Eastern Europe. *Social Forces*, 86(1), 231–255. https://doi.org/https://doi.org/10.1353/sof.2007.0087

2) Aspinall, P.J. (1997). The conceptual basis of ethnic group terminology and classifications. *Social Science & Medicine*, *45*(5), 689–698. https://doi.org/10.1016/S0277-9536(96)00386-3

3) Bennett, T. (1997). "Racial" and ethnic classification: Two steps forward and one step back? *Public Health Reports*, 112(6), 477–480. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1381925/

4) Bonnett, A., & Carrington, B. (2000). Fitting into categories or falling between them? Rethinking ethnic classification. *British Journal of Sociology of Education,* 21(4), 487–500. https://doi.org/10.1080/713655371

5) Canut, C. (2019). Tell Me That I Am Not a *Ciganin*, Damn Your Mother! The Social and Political Consequences of Enregisterment in Bulgaria. *Signals and Society*, *7*(3), 398–426. https://doi.org/10.1086/704985

6) Felouzis, G. (2010). Ethnic Categories in Sociology: Material for a Debate. *Revue Française de Sociologie, Suppl. An Annual English Selection*, 51, 145–150. https://www.proquest.com/scholarly-journals/ethnic-categories-sociology-material-debate/docview/670679089/se-2?accountid=15133

7) Foster, B.L., & White, G.M. (1982). Ethnic-Identity and Perceived Distance between Ethnic Categories. *Himan Organization*, *41*(2), 121–130. https://doi.org/10.17730/humo.41.2.24xpnu1187574543

8) Kaufman, J.S., & Cooper, R.S. (2001). Commentary: Considerations for use of racial/ethnic classification in etiologic research. *American Journal of Epidemiology*, 154(4), 291–298. https://doi.org/10.1093/aje/154.4.291

9) Kisfalusi, D. (2018). Ethnic classification among secondary school teachers and students in Hungary. *Corvinus Journal of Sociology and Social Policy*, *9*(1), 35-54. https://doi.org/https://doi.org/10.14267/CJSSP.2018.1.02

10) Lachowsky, N. J., Saxton, P. J. W., Dickson, N. P., Hughes, A. J., Jones, R. G., Clark, T. C., Ho, E., Summerlee, A. J. S., & Dewey, C. E. (2020). Ethnicity classification systems for public health surveys: experiences from HIV behavioural surveillance among men who have sex with men. *BMC Public Health*, *20*, 1–13. https://doi.org/https://doi.org/10.1186/s12889-020-09517-4

11) Lauderdale, D. S., & Goldberg, J. (1996). The expanded racial and ethnic codes in the Medicare data files: their completeness of coverage and accuracy. *American journal of public health*, *86*(5), 712–716. https://doi.org/10.2105/ajph.86.5.712.

12) Lee, S., Satter, D.E., & Ponce, N.A. (2009). Effect of Race and Ethnicity and ethnic theory classification on survey estimates: Anomaly of the Weighted Totals of American Indians and Alaska Natives. *American Indian and Alaska Native Mental Health Research*, *16*(3), 1–15.

13) Lewis, C., Cohen, P. R., Devyani, B., Levine, E. M., & Waseem, K. (2023). Race and Ethnic Categories: A Brief Review of Global Terms and Nomenclature. *Cureus*, *15*(7), e41253. https://doi.org/10.7759/cureus.41253

14) Marini, M., Waterman, P.D., Breedlove, E.R., Chen, J.T., Testa, C., Pardee, D.J., LeBlanc, M., Reisner, S.L., Oendari, A., & Krieger, N. (2023). Using Implicit Measures of Discrimination: White, Black, and Hispanic Participants Respond Differently to Group-Specific Racial/Ethnic Categories vs. the General Category "People of Color" in the USA. *Journal of Racial and Ethnic Health Disparities*, *10*(4), 1682–1692. https://doi.org/10.1007/s40615-022-01353-z

15) Mateos, P. (2007). A review of name-based ethnicity classification methods and their potential in population studies. *Population Space and Place*, *13*(4), 243–263. https://doi.org/10.1002/psp.457

16) Morning, A. (2008). Ethnic classification in global perspective: A cross-national survey of the 2000 census round. *Population Research and Policy Review*, *27*(2), 239–272. https://doi.org/10.1007/s11113-007-9062-5

17) Ringelheim, J. (2011). Ethnic categories and European human rights law. *Ethnic and Racial Studies*, *34*(10), 1682–1696. https://doi.org/10.1080/01419870.2010.542249

18) Roman, R.B. (2018). "Neither here, nor there": Belonging, ambiguity, and the struggle for recognition among "in-between" Finnish Kaale. *Romani Studies*, *28*(2), 239–262. https://doi.org/10.3828/rs.2018.10

19) Ruane, J. & Todd, J. (2004). The Roots of Intense Ethnic Conflict may not in fact be Ethnic: Categories, Communities and Path Dependence: European Journal of Sociology. *Archives Européennes de Sociologie*, *45*(2), 209–232. https://doi.org/https://doi.org/10.1017/S0003975604001432

20) Salo, M.T. (1979). Gypsy Ethnicity: Implications of Native Categories and Interaction for Ethnic Classification. *Ethnicity*, *6*(1), 73–96.

21) Salway, S. (2018). Working "with and against" ethnic categories to understand and tackle health inequalities: illustrations from research and practice relating to infant mortality in England. *European Journal of Public Health*, *28*, 30–31. https://doi.org/10.1093/eurpub/cky047.036

22) Schaeffer, M. (2013). Which groups are mostly responsible for problems in your neighbourhood? The use of ethnic categories in Germany. *Ethnic and Racial Studies*, *36*(1), 156–178. https://doi.org/10.1080/01419870.2011.644311

23) Sigona, N. (2003). How can a "nomad" be a "refugee"? Kosovo Roma and labelling policy in Italy. *Sociology*, *37*(1), 69–79. https://doi.org/10.1177/0038038503037001445

24) Sookrajowa, S.S. (2021). Legibility and the Politics of Ethnic Classification of the Population in the National Census of Mauritius: A Statist Perspective. *Nationalism and Ethnic Politics*, 27(2), 128–148. https://doi.org/10.1080/13537113.2021.1888449

25) Surdu, M. (2019). Why the "real" numbers on Roma are fictitious: Revisiting practices of ethnic quantification. *Ethnicities*, 19(3), 486–502. https://doi.org/10.1177/1468796819833424

26) Wolf, S.M. (2006). Debating the Use of Racial and Ethnic Categories in Research. *The Journal of Law, Medicine & Ethics*, 34(3), 483–486. https://doi.org/https://doi.org/10.1111/j.1748-720X.2006.00059.x

27) Zhang, F., & Finkelstein, J. (2019). Inconsistency in race and ethnic classification in pharmacogenetics studies and its potential clinical implications. *Pharmacogenomics and Personalized Medicine*, 12: 107–123. https://doi.org/10.2147/PGPM.S207449

Appendix 3. Sources identified by AG members

1) The Traveller Movement (No Date). *Romani (Gypsy), Roma and Irish Traveller History and Culture.* Available at: https://travellermovement.org.uk/gypsy-roma-and-traveller-history-and-culture

2) Annan-Callcott, G. (No Date). *The problem with ethnicity categories in UK health data*. Available at: https://wellcome.org/news/ethnicity-categories-uk-health-data

3) ONS (2022). *Gypsies' and Travellers' lived experiences, health, England and Wales*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthcaresystem/bulletins/gypsiesandtravellerslivedexperienceshealthenglandandwales/2022

4) ONS (2024). *Quality of ethnicity data in health-related administrative data sources by sociodemographic characteristics, England Articles*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthinequalities/articles/understandingconsistencyofethnicitydatarecordedinhealthrelatedadministrativedatasetsinengland2011to2021/previousReleases

5) Lipott, S. (2012). The Roma as a Protected Minority? Policies and Best Practices in the EU. *Romanian Journal of European Affairs*, 12(4): 78-98. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2233866

6) Scottish Government (2016). *Scottish Social Attitudes 2015, Attitudes to Discrimination and Positive Action*. Available at: https://www.gov.scot/publications/scottish-social-attitudes-2015-attitudes-discrimination-positive-action/

7) Shelter Scotland (2015). *Gypsy/Travellers and Financial Exclusion: Examination of Best Practice in the development of Financial Capability*. Available at: https://scotland.shelter.org.uk/professional_resources/policy_library/gypsytravellers_and_financial_exclusion_-_an_examination_of_best_practice_in_the_development_of_financial_capability

8) Dowd, L. (2022). *Death threats are a way of life': Gypsies and Irish Travellers and Muslims 'least-liked' in UK, survey finds*. Available at: https://news.sky.com/story/death-threats-are-a-way-of-life-gypsies-and-irish-travellers-and-muslims-least-liked-in-uk-survey-finds-12524098

9) Wales Government (2018). *Enabling Gypsy, Roma & Travellers*. Available at: www.gov.wales/sites/default/files/publications/2019-02/enabling-gypsies-roma-and-travellers_0.pdf

10) Wales Government (2022). *Anti-racist Wales Action Plan: 2022*. Available at: https://www.gov.wales/anti-racist-wales-action-plan-2022

11) Council of Europe (2020). *Strategic Action Plan for Roma and Traveller Inclusion (2020-2025)*. Available at: https://rm.coe.int/coe-strategic-action-plan-for-roma-and-traveller-inclusion-en/16809fe0d0

12) Equality and Human Rights Commission (2023). *Equality and Human Rights Monitor: Is Wales Fairer?* Available at: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.equalityhumanrights.com/

sites/default/files/2023/Is%20Wales%20Fairer%20Equality%20and%20Human%20Rights%20Monitor-%20English-%20accessible%20PDF.pdf

13) ONS (2023). *Analytical Learning Points for Ethnicity Data in health administrative data sources*. Available at: https://analysisfunction.civilservice.gov.uk/policy-store/analytical-learning-points-for-ethnicity-data/

AG members:

- Jackie Bolton, BA (Hons.) – Traveller
- Joanna Castle - Data and Analytical Lead for Inclusion Health Populations at Office for Health Improvement and Disparities (Department of Health and Social Care)
- Rosa Cisneros, PhD – Spanish Roma
- Stacy Hodgkins, BA (Hons.) – English Romany
- Allison Hulmes, MSc – Welsh Romany
- Shamus McPhee, MA, PG Dip. – Scottish Gypsy-Traveller (Nacken)
- Marta Pineda Moncusí, MSc, PhD - Postdoctoral Researcher in Health Data, Oxford University
- Emma Morgan – Policy & Engagement Manager at Understanding Patient Data (NHS Confederation)
- Josie O'Driscoll, CEO of GATE Herts – Irish Traveller
- Cameron Razieh, Research Fellow in Epidemiology, Leicester Real World Evidence Unit, University of Leicester
- Rebecca Asher and Emily Jesper-Mir – Partnerships and Engagement Managers in Data for Science and Health Team at Wellcome Trust

Appendix 4. Currently used ethnic category tables in England, Scotland, Wales (Relevant categories highlighted)

1) Census 2021 ethnic categories (accessed 13/12/2024)

1. White: English/Welsh/Scottish/Northern Irish/British
2. White: Irish
3. White: Gypsy or Irish Traveller
4. White: Roma
5. White: Other White
6. Mixed/multiple ethnic groups: White and Black Caribbean
7. Mixed/multiple ethnic groups: White and Black African
8. Mixed/multiple ethnic groups: White and Asian
9. Mixed/multiple ethnic groups: Other Mixed
10. Asian/Asian British: Indian
11. Asian/Asian British: Pakistani
12. Asian/Asian British: Bangladeshi
13. Asian/Asian British: Chinese
14. Asian/Asian British: Other Asian
15. Black/African/Caribbean/Black British: African
16. Black/African/Caribbean/Black British: Caribbean
17. Black/African/Caribbean/Black British: Other Black
18. Other ethnic group: Arab
19. Other ethnic group: Any other ethnic group

2) GPES (General Practitioner Extraction Service) ethnic categories (accessed 13/12/2024)

1. British
2. Irish
3. Traveller
4. Any other White background
5. White and Black Caribbean
6. White and Black African
7. White and Asian
8. Any other Mixed background

9. Indian
10. Pakistani
11. Bangladeshi
12. Chinese
13. Any other Asian background
14. African
15. Caribbean
16. Any other Black background
17. Arab
18. Any other ethnic group

3) HES (Hospital Episode Statistics) ethnic categories (accessed 13/12/2024)

1. British (White)
2. Irish (White)
3. Any other White background
4. White and Black Caribbean (Mixed)
5. White and Black African (Mixed)
6. White and Asian (Mixed)
7. Any other Mixed background
8. Indian (Asian or Asian British)
9. Pakistani (Asian or Asian British)
10. Bangladeshi (Asian or Asian British)
11. Chinese (Other ethnic group)
12. Any other Asian background
13. African (Black or Black British)
14. Caribbean (Black or Black British)
15. Any other Black background
16. Any other ethnic group

4) Public Health Scotland – Ethnic categories (accessed 13/12/2024)

1. Scottish
2. Other British

3. Irish
4. Gypsy/Traveller
5. Polish
6. Roma
7. Showman/Showwoman
8. Other with ethnic group
9. Any mixed or multiple ethnic groups
10. Pakistani, Scottish Pakistani or British Pakistani
11. Indian, Scottish Indian or British Indian
12. Bangladeshi, Scottish Bangladeshi or British Bangladeshi
13. Chinese, Scottish Chinese or British Chinese
14. Other Asian, Scottish Asian or British Asian
15. African, Scottish African or British African
16. Caribbean or Black
17. Arab, Scottish Arab or British Arab
18. Other ethnic group

5) Public Health Wales – Ethnic categories (accessed 13/12/2024)

1. Any White Background, including Welsh, English, Scottish, Northern Irish, Irish, British
2. Gypsy or Irish Traveller
3. White and Black Caribbean
4. White and Black African
5. White and Asian
6. Any other mixed background / multiple ethnic background
7. Indian
8. Pakistani
9. Bangladeshi
10. Chinese
11. Any other Asian background
12. Caribbean
13. African
14. Any other Black background

15. Arab

16. Any other ethnic group

6) Ethnic categories in Health and Social Care in Northern Ireland (accessed 13/12/2024)

1. White

2. Chinese

3. Irish Traveller

4. Indian

5. Pakistani

6. Bangladeshi

7. Black Caribbean

8. Black African

9. Black other

10. Mixed ethnic group

11. Any other ethnic group