Is Your Ethnicity Data Up To Standard?

Alana B. McCambridge *PhD*
Person Centred Rehabilitation Research Centre, Auckland University of Technology, Auckland, New Zealand

McCambridge, A. B. (2025). Is your ethnicity data up to standard? *New Zealand Journal of Physiotherapy, 53*(3), 101–103. https://doi.org/10.15619/nzjp.v53i3.526

Key Words: Data Integrity, Editorial, Ethnicity

Ethnicity profoundly shapes an individual's experiences and health outcomes and is a robust evidence-based marker of need (Harris et al., 2024; Loring et al., 2024). In New Zealand, research on ethnicity accuracy between self-reported ethnicity and secondary National Health Index (NHI) and primary healthcare records found that ethnicity data were inaccurate for 9.5–11% of individuals (Blackmore et al., 2024). The accuracy of ethnicity data was significantly worse for Māori (Blackmore et al., 2024). Ethnicity is an important variable to consider in both clinical and research contexts. In physiotherapy, ethnic disparities have been demonstrated across multiple settings, with worse outcomes consistently reported for marginalised ethnic groups (Patel et al., 2024). Yet the rigour with which ethnicity is measured, analysed, and reported appears to be lacking. One reason for this may be that people working with health data, such as clinicians and researchers, could be unaware of the standards that exist for ensuring ethnicity data is robust and consistent.

ETHNICITY AS A VARIABLE

In New Zealand, there are two key documents that guide our understanding of ethnicity and how ethnicity data should be standardised across the health sector. First, the 'Statistical standard for ethnicity V1.0.0' (Stats NZ, 2023) sets the conceptual definition of ethnicity, defining ethnicity as "… a measure of cultural affiliation, as opposed to race, ancestry, nationality or citizenship. Ethnicity is self-perceived and people can belong to more than one ethnic group". This document promotes the standardisation of ethnicity to ensure data consistency and comparability in all official government statistics. In alignment with this are "The Ethnicity Data Protocols for the Health and Disability Sector" (Ministry of Health, 2017). This document outlines the minimum standards that apply across the health and disability sector to guide the standardised collection, recording, and output of ethnicity data. Unfortunately, awareness and commitment to these protocols appears to be lacking, which may be undermining our ability to monitor ethnic disparities, track progress, and respond effectively to population needs.

To ensure the standardisation of ethnicity data in your research or clinical practice, it is important to be familiar with the key elements that underpin the concept of ethnicity as well as requirements specified by the data protocols.

KEY ELEMENTS OF ETHNICITY

There are three key elements to comprehend when conceptualising ethnicity. First, ethnicity must be self-identifiable. This means individuals must be able to define and express their own ethnic affiliations, rather than having them assigned or interpreted by others, unless they are incapable of doing so (e.g., infants). Research on ethnicity data collection in New Zealand GP settings, has shown that some receptionists feel discomfort when asking for a person to self-identify their ethnicity (Neuwelt et al., 2014). More concerningly, some believed they could make a valid assumption about a person's ethnicity (Neuwelt et al., 2014), which is not appropriate and would degrade the validity of the data. However, with information and training about ethnicity data quality, staff are more motivated to improve how they collect ethnicity data in their practices (Neuwelt et al., 2014).

Second, individuals may identify with multiple ethnicities. This means any method of collecting ethnicity data that restricts individuals to selecting only one ethnic group is inaccurate and misrepresents the concept of ethnicity. Research on ethnicity accuracy in secondary NHI and primary care datasets found that those who identify with multiple ethnicities were 95% less likely to have an accurate ethnicity record compared to those who identify with a single ethnicity (Blackmore et al., 2024). In the latest census, 15.2% of the New Zealand population identified with more than one ethnicity (Stats NZ, 2025), therefore the impact of such errors is potentially very large.

Third, an individual's ethnicity can change. This is termed ethnic mobility and it means individuals must be given the opportunity to update and change their ethnicity. The ethnicity protocols require that individuals are asked about their ethnicity at least once every three years, and that healthcare providers offer an opportunity to update a patient's self-identified ethnicity at each interaction (Ministry of Health, 2017).

A BRIEF OVERVIEW OF HOW TO COLLECT, COLLATE, AND REPORT ETHNICITY DATA

As mentioned, the "Ethnicity Data Protocols" guide the collection, classification, and output of ethnicity data in the health and disability sector (Ministry of Health, 2017). Inaccuracy of ethnicity data could occur by failing to meet the requirements set out by the data protocols at any of these stages. Anyone who is responsible for collecting or interpreting ethnicity data should engage with the "Ethnicity Data Protocols" document for a comprehensive understanding; however, several key points related to the requirements are summarised below.

Collection of ethnicities

It is not appropriate to simply ask "What is your ethnicity?" To ensure appropriate collection of ethnicity data, it is a requirement that the standardised ethnicity question is used (Ministry of Health, 2017). The paper form version of this question, including the way in which it is formatted, should be reproduced using the actual standardised graphic or abiding by the minimum requirements to replicate the question (Ministry of Health, 2017). Guidance also exists on how to appropriately collect ethnicity data electronically, verbally using a script (i.e., "I am going to read out a list of ethnic groups. Can you tell me which ethnic group or groups you belong to…?"), and as a proxy when the individual is unable to answer the question themself. All approaches to collecting ethnicity data must ensure individuals are able to identify with multiple ethnicities.

Grouping of ethnicities

In New Zealand, a four-tiered classification structure exists to aggregate ethnicities into sensible groups. The Level 1 codes are broad groupings (e.g., Māori, Pacific peoples, Asian, Middle Eastern/Latin American/African, European, and Other), whereas the level 4 codes are the most granular. For example, someone who is Thai will be classified as 'Thai' at level 4, 'Other South East Asian' at level 3, 'Southeast Asian' at level 2, and 'Asian' at level 1. Previous iterations of the data protocols required ethnicity be collected at level 2; however, the current version of the protocols requires ethnicity to be collected at level 4 (highest detail). It is therefore inconsistent with the data protocols to collect ethnicity data at level 1 (least detailed).

Data should be collected at level 4 to ensure the greatest granularity of the data, such that the data are sufficiently detailed to best reflect a person's ethnic affiliation and experience (Gurney et al., 2018). When grouping ethnicity data, it is important to consider what is being made invisible in your analysis with such high-level groupings. For example, a study on cardiovascular disease in New Zealand found that the prevalence of health failure is 1.7 times higher in Cook Island Māori in comparison to Niueans, and 1.4 times higher than Tongans and Samoans (Winter-Smith et al., 2021). These findings have important implications for the management of cardiovascular disease in Pacific peoples and would not have been visible if the analysis did not investigate each Pacific ethnicity separately.

Output methods

There are three different output methods of ethnicity data that can be used: total response (overlapping), prioritised, and sole/combination. Each method has distinct considerations that users should be cognisant of.

New Zealand government data are required to use total response (overlapping), as this method ensures all individuals are counted in each of the ethnic groups they are affiliated with. Therefore, this output method aligns well with the conceptual definition of ethnicity, ensuring multiple ethnicities are permissible. However, this output method may pose analytical challenges given an individual who identifies with more than one ethnicity will be counted in more than one group and the sum of each group will therefore be greater than the total population.

In the health and disability sector, prioritised ethnicity is the most common output method. The prioritised method assigns an individual to a single ethnicity based on a predefined prioritisation order. This is somewhat in contrast to the conceptual definition of ethnicity, which states that an individual can self-identify with any and all ethnicities of their choosing. Assigning an individual to a single ethnic category may misrepresent the complexity of their identity, especially when they identify with multiple ethnic groups or when the assigned ethnicity does not reflect the one they most strongly associate with. Further considerations with this output method are that the prioritisation order, although constructed by policy importance, are essentially arbitrary thresholds (e.g., level 2 prioritisation ranks Fijian above Tongan, Indian above African), and this method over-represents some groups at the expense of other groups (McCambridge, 2024). For example, those who identify as Māori and Samoan would be represented in the data as Māori only and not counted in the Pacific data.

Finally, sole/combination output is another possible output method. It creates categories of single ethnicities and combination ethnicities. For example, there may be groups of people who identify as Māori only, European only, and those who identify as Māori/European combined. This method is relatively uncommon to use in health and would likely require large sample sizes for each of the sole and combination groupings.

As explained above, each output method has implications that should be taken into consideration when analysing and interpreting the data. As a minimum requirement, the output method that is used in the handling of ethnicity should be clearly reported.

FINAL THOUGHTS

Ethnicity is the strongest, evidence-based marker of health need we have in New Zealand (Loring et al., 2024). Ethnicity data are commonly encountered by clinicians and researchers in both clinical practice or research settings. Yet deep engagement with what ethnicity actually means, and why it matters to have high quality ethnicity remains underdeveloped, despite its central role in addressing health inequities. Regular audits and ongoing education around the minimum standards for ethnicity data collection, analysis, and reporting are essential until consistent, widespread adoption is achieved across both clinical and research contexts.

ADDRESS FOR CORRESPONDENCE

Alana McCambridge, Person Centred Rehabilitation Research Centre, Auckland University of Technology, Auckland, New Zealand.

Email: alana.mccambridge@aut.ac.nz

REFERENCES

Blackmore, B., Elston, M., Loring, B., Reid, P., & Tamatea. J. (2024). Accuracy of ethnicity records at primary and secondary healthcare services in Waikato region, Aotearoa New Zealand. *New Zealand Medical Journal, 137*(1602), 111–124. https://doi.org/10.26635/6965.6445

Gurney, J. K., Scott, N., Thompson, G., Turner, S., Anson, J., Cragg, M., Doherty, J., Wall, M., Moke, T., Blundell, R., & Coote, P; Hei Āhuru Mōwai – National Māori Cancer Leadership Group. (2018). New Zealand's revised ethnicity data protocols must not become a shelved document: A challenge from Hei Ahuru Mowai [letter]. *New Zealand Medical Journal*, *131*(1470), 104–106. https://nzmj.org.nz/media/pages/journal/vol-131-no-1470/new-zealands-revised-ethnicity-data-protocols-must-not-become-a-shelved-document-a-challenge-from-hei-ahuru-mowai/7cfa8cc86c-1696472495/new-zealands-revised-ethnicity-data-protocols-must-not-become-a-shelved-document-a-challenge-from-hei-ahuru-mowai.pdf

Harris, R., Cormack, D., Waa, A., Edwards, R., & Stanley, J. (2024). The impact of racism on subsequent healthcare use and experiences for adult New Zealanders: A prospective cohort study. *BMC Public Health*, *24*(1), Article 136. https://doi.org/10.1186/s12889-023-17603-6

Loring, B., Reid, P., Curtis, E., McLeod, M., Harris, R., & Jones, R. (2024). Ethnicity is an evidence-based marker of need (and targeting services is good medical practice). *New Zealand Medical Journal, 137*(1603), 9–13. https://doi.org/10.26635/6965.e1603

McCambridge, A. B. (2024). The need for transparent reporting of ethnicity in health research [letter]. *New Zealand Medical Journal, 137*(1603), 147–149. https://nzmj.org.nz/media/pages/journal/vol-137-no-1603/the-need-for-transparent-reporting-of-ethnicity-in-health-research/4a01320053-1727148407/6686-and-response.pdf

Ministry of Health. (2017). *Ethnicity data protocols. Version 1.1.* [HISO 10001:2017]. https://www.tewhatuora.govt.nz/assets/Our-health-system/Digital-health/Health-information-standards/HISO-10001-2017-Ethnicity-Data-Protocols.pdf

Neuwelt, P., Crengle, S., Cormack, D., McLeod, M., & Bramley, D. (2014). General practice ethnicity data: Evaluation of a tool. *Journal of Primary Health Care, 6*(1), 49–55.

Patel, R. M., Prusynski, R. A., Williams-York, B., Matthews, N. D., & VanHoose, L. (2024). Disparities in physical therapy outcomes based on race and ethnicity: A scoping review. *Physical Therapy, 104*(10), Article pzae090. https://doi.org/10.1093/ptj/pzae090

Stats NZ. (2023). *Statistical standard for ethnicity* V1.0.0. https://aria.stats.govt.nz/aria/#StandardView:uri=http://stats.govt.nz/cms/StatisticalStandard/vv0ovwUoTSSVDhpt

Stats NZ. (2025, 29 July). *Place and ethnic group summaries*. https://www.stats.govt.nz/tools/place-and-ethnic-group-summaries/

Winter-Smith, J., Selak, V., Harwood, M., Ameratunga, S., & Grey, C. (2021). Cardiovascular disease and its management among Pacific people: A systematic review by ethnicity and place of birth. *BMC Cardiovascular Disorders, 21*(1), Article 515. https://doi.org/10.1186/s12872-021-02313-x