

VIEW Ethnicity Protocol

Ethnicity is assigned to an individual based on a prioritisation output. The prioritisation ethnicity protocol adopted by VIEW is based on the Statistics New Zealand ethnicity prioritisation method, and is the most frequently used output method in Ministry of Health statistics. The table below shows level 2 ethnicity codes and their corresponding priority. More information on prioritised output can be found in Appendix A

Table 1

Level 2 ethnic codes

Ethnic Group code	Ethnic Group code description	Ethnic Group priority	Revised VIEW priority
10	European not further defined	21	
11	NZ European	22	
12	Other European	20	
21	NZ Maori	1	
30	Pacific Island not further defined	9	
31	Samoan	7	
32	Cook Island Maori	6	
33	Tongan	5	
34	Niuean	4	
35	Tokelauan	2	
36	Fijian	3	
37	Other Pacific Island	8	
40	Asian not further defined	14	
41	Southeast Asian	10	12
42	Chinese	12	11
43	Indian	11	10
44	Other Asian	13	
51	Middle Eastern	17	
52	Latin American / Hispanic	15	
53	African	16	
54	Other (retired on 1/07/2009)	19	
61	Other ethnicity	18	
94	Don't know	94	
95	Refused to answer	95	
97	Response unidentifiable	97	
99	Not stated	99	

PREDICT 2015 baseline data – Unique ethnicity codes

Ethnicity data used in VIEW comes from two sources – PREDICT and Ministry of Health. When patients are enrolled into PREDICT, their ethnicity are recorded across three ethnicity inputs fields (allowing for the self-identification of up to 3 ethnicity responses). In addition, the Ministry of Health has provided us with a 2015 update of the NHI Demographic Lookup table, containing the demographic data for 7.7 million unique eNHI. Similarly, up to three ethnicity codes are provided (allowing for the self-identification of up to three ethnicity responses). In total, each patient has up to 6 codes that represent their ethnicity.

All unique responses provided from each of the ethnicity fields in the PREDICT 2015 Baseline Data

Source	Variable name	Ethnicity Codes
PREDICT 2015	pt_ethnic_group_1	10 11 12 21 30 31 32 33 34 35 36 37 40 41 42 43 44 51 52 53 54 441 442 443 444 44411 44412 44413 44414 44415 NA
	pt_ethnic_group_2	10 11 12 21 30 31 32 33 34 35 36 37 40 41 42 43 44 51 52 53 54 99 441 443 44411 44412 44414 NA
	pt_ethnic_group_3	10 11 12 21 30 31 32 33 34 35 36 37 40 41 42 43 44 51 52 53 54 99 441 44411 44414 NA
Ministry of Health 2015	nhi_ethnicg1	10 11 12 21 30 31 32 33 34 35 36 37 40 41 42 43 44 51 52 53 54 61 94 95 97 99
	nhi_ethnicg2	10 11 12 21 30 31 32 33 34 35 36 37 40 41 42 43 44 51 52 53 54 61 94 95 97 99 NA
	nhi_ethnicg3	10 11 12 21 30 31 32 33 34 35 36 37 40 41 42 43 44 51 52 53 61 97 99 NA

NB: There are no NAs in “nhi_ethnicg1”

Procedure for Ethnicity Allocation

The procedure assigns one single ethnicity to each individual. The ethnicity response (there are 6 in total) of each individual is read by the programme using the prioritisation protocol. The programme checks each of the 6 ethnicity fields of a person, and determines which single ethnicity will be assigned. The programme checks each row of data and executes the following command in this order:

- 1) Is this person Maori? If yes, write "NZMaori", otherwise next question.
- 2) Is this person Pacific? If yes, write "Pacific", otherwise next question.
- 3) Is this person Indian? If yes, write "Indian", otherwise next question.
- 4) Is this person Chinese? If yes, write "Chinese", otherwise next question.
- 5) Is this person Asian? If yes, write "Asian", otherwise next question.
- 6) Is this person MELAA? If yes, write "MELAA", otherwise next question.
- 7) Is this person Other? If yes, write "Other", otherwise next question.
- 8) Is this person European? If yes, write "European", otherwise next question.
- 9) Is the ethnicity unknown, not answered, not identifiable? If yes, write "No_not_stated".

NB: MELAA = Middle Eastern, Latin American, African

VIEW REVISED Procedure for Ethnicity Allocation

- 1) Is this person Maori? If yes, write "NZMaori", otherwise next question.
- 2) Is this person Pacific? If yes, write "Pacific", otherwise next question.
- 3) Is this person Indian? If yes, write "Indian", otherwise next question.
- 4) Is this person Chinese? If yes, write "Chinese", otherwise next question.
- 5) Is this person Asian? If yes, write "Asian", otherwise next question.
- 6) Is this person European? If yes, write "European", otherwise next question.
- 7) Is this person MELAA? If yes, write "MELAA", otherwise next question.
- 8) Is this person Other? If yes, write "Other", otherwise next question.
- 9) Is the ethnicity unknown, not answered, not identifiable? If yes, write "No_not_stated".

Multiple Ethnicities

Any individuals with multiple ethnicity responses will be assigned the higher priority of ethnicity.

Example 1 – If a patient is recorded as Maori (21) and Samoan (31), then they are recorded as “Maori”. This is because the programme asks whether this person is “Maori” first. With the answer being yes, “Maori” is recorded. The programme then moves onto the next person instead of asking whether or not they are Pacific.

Example 2 – If a person is recorded as Chinese (42), Southeast Asian (41), and NZ European (11), then they are recorded as Chinese. With Chinese being the highest priority, the person is assigned “Chinese” and the programme moves onto the next person.

NB: “Asian” contains Southeast Asian (41) which has a higher priority compared to Indian and Chinese (see Table 1). However, due to its relatively small population, the Southeast Asian group will be included in the “Asian” group, and thus not prioritised over Indian or Chinese. This is the ONLY exception to the prioritisation order!

The use of “OTHER” Ethnicity

This classification should be clearly defined. The term “Other” does in fact have its own ethnicity coding. It should not be used as a category for which miscellaneous or small populations are assigned as a matter of convenience. Previously, Middle Eastern (51), Latin American/Hispanic (52), and African (53), were frequently included in the OTHER ethnic group. Since 2009 (I think), Statistics New Zealand and the MOH have adopted a new category called MELAA which incorporates codes 51-53. A distinction between MELAA and Other is therefore created. There are two codes (and there should only be two codes), for Other Ethnicity – 54 (pre-2009) and 61 (post-2009).

Original “ag_eth” Classification

Label	Code
Maori	21
Pacific	30, 31, 32, 33, 34, 35, 36, 37
Indian	43, (36 & 43)
Asian	40, 41, 42, 44, 441, 442, 443, 444, 44411, 44412, 44414
Other	51, 52, 53, 54
European	10, 11, 12, 94, 95, 96, 99, " ", ""

Problems with above coding convention:

- “44415” is missing from Asian group
- MELAA codes (51-53) are recorded as “Other Ethnicity”
- “Other Ethnicity” code (61) missing
- European group contains residual codes (94, 95, 96, 99, " ", "")
- “Chinese” are not represented clearly

Distribution of original “ag_eth” (all unique individuals at baseline)

Frequency

Asian	European	Indian	NZMaori	Other	Pacific	<NA>
45308	276933	39205	62181	8907	59305	306

NB: There should be no NA values since **nhi_ethnicg1** contains no NAs

Proportion

Asian	European	Indian	NZMaori	Other	Pacific	<NA>
0.092	0.563	0.080	0.126	0.018	0.121	0.001

NEW “view_ag_eth” Classification

Label	Code
Maori	21
Pacific	30, 31, 32, 33, 34, 35, 36, 37
Indian	43, (36 & 43)
Chinese	42
Asian	40, 41, 44, 441, 442, 443, 444, 44411, 44412, 44414, 44415
MELAA	51, 52, 53
Other	54, 61
European	10, 11, 12
No_not_stated	94, 95, 96, 99, " ", ""

“Other” includes individuals who write “Klingon” or “Martian” as their response.

This list of ethnic groups can be combined as suited to the individual study, however the default coding for VIEW should be that “MELAA” and “Other” will be combined into “Other”. As this is a very heterogeneous group, it may be left out of analyses that focus on ethnic-specific analyses.

“No_not_stated” is defined rather than the default “NA”. The reason is that the MOH have codes precisely for these situation, ranging from “Don’t know” (94), “Refused to Answer” (95), to “Not Stated” (99). If you’re reporting the status of everyone in your cohort of interest, this should be stated as being missing data on ethnicity and not combined with “Other”, as they represent two different types of data.

In previous merges, the European group included “Other” and “NA”. The new coding allows European to be more clearly defined.

Distribution of proposed new “ag_eth2” (all unique individuals at baseline)

Frequency

Asian	Chinese	European	Indian	MELAA	No_not_stated	NZMaori
18745	26563	276433	39205	6797	654	62181
Other	Pacific	<NA>				
2262	59305	0				

Proportion

Asian	Chinese	European	Indian	MELAA	No_not_stated	NZMaori
0.038	0.054	0.562	0.080	0.014	0.001	0.126
Other	Pacific	<NA>				
0.005	0.121	0.000				

Appendix A

Prioritisation Output for Ethnicity

In prioritised output, each respondent is allocated to a single ethnic group using the priority system (Māori, Pacific peoples, Asian, other groups except NZ European; and NZ European). The aim of prioritisation is to ensure that where some need exists to assign people to a single ethnic group, ethnic groups of policy importance, or of small size, are not swamped by the NZ European ethnic group.

This output type is the one most frequently used in Ministry of Health statistics and is also widely used in the health and disability sector for funding calculations, monitoring changes in the ethnic composition of service utilisation, and so on. Its advantage is that it produces data that are easy to work with as each individual appears only once so the sum of the ethnic group populations will add up to the total New Zealand population.

When ethnicity data is to be output to the Ministry of Health National Systems and more than three ethnicities are available to send, the prioritisation method described in the protocols must be used. This will ensure consistency within the national collections.

Limitations are that prioritised output:

- places people in specific (high priority because of policy importance) ethnic groups which simplifies yet biases the resulting statistics
- over-represents some groups at the expense of others – for example, Māori gain at the expense of Pacific peoples (approximately 31,542) and Pacific peoples gain at the expense of other groups (34,602) of which most are Pacific/European (30,018)
- goes against the principle of self-identification.

One of the main criteria stipulated in the definition of ethnicity is that a person can belong to more than one ethnic group. The ethnicity question caters for multiple responses. However, the question does not ask people to indicate the ethnic group with which they identify the most strongly; instead, prioritisation makes this choice for them. The question is to remain the same for the 2006 census so, to ensure numerator and denominator consistency (see Section 1.5), asking people to state the ethnicity with which they identify the 'most strongly' is not an option.