

GLOSSARY

Glossary of terms relating to ethnicity and race: for reflection and debate

R Bhopal

J Epidemiol Community Health 2004;**58**:441–445. doi: 10.1136/jech.2003.013466

This glossary focuses on the concepts and terminology used in the study of the health of minority ethnic and racial groups. It is hoped that it will stimulate debate on this subject so that an internationally applicable glossary may emerge.

In rising to the editor's challenge to write this glossary I am conscious that I am summarising and reigniting longstanding controversies. This paper comprises a personal viewpoint, a summary of the current position, and a signpost for the way forward. The glossary is limited by my knowledge and experience, which mainly derives from the UK population context, with some observations in the USA and Europe. It is unlikely, however, that the "state of the art" is advanced enough to achieve more than a context-specific perspective. In the future, if conceptual advances can be achieved, and principles on how to operationalise the concepts agreed, an internationally applicable glossary may emerge.

As a Punjabi born Indian raised in Scotland I found the popular UK label "Asian" to describe people like me as simplistic. In 1984 I learned that this label was embedded in the scientific literature. In publications in the 1980s I took the unusual step of defining Asian, for example, "For the purposes of this study, Asian refers to persons whose ancestry is from the Indian subcontinent". I also tried italicising the word Asian and putting it in quotations to alert the reader to the limited use of the word. In retrospect these steps were insufficient. I realised that in the United States Asian was interpreted as far Eastern Asian populations.¹ In my writing I started to provide a statement on my use of terms, sometimes as an introductory paragraph as indicated in the appendix.

This paper shares my struggle, with the hope that others will help resolve the problems. In 1990, I wrote that an internationally agreed vocabulary was the ideal.² Sadly, there has been little progress towards this goal. This paper will also end with a plea for work on an internationally agreed glossary. The full bibliography on the journal web site (<http://www.jech.com/supplemental>) gives a sample of the scientific debate.

NEED FOR THE VARIABLES OF ETHNICITY AND RACE IN EPIDEMIOLOGY AND PUBLIC HEALTH

The concepts of ethnicity and race in health care and public health raise difficult ethical issues,

which have seldom been explicitly considered.³ Ethnicity and race are controversial variables in epidemiology and public health, including the many branches of these disciplines, and yet they are of central interest.^{4–8} The fuel of epidemiology is the analysis of differences in the pattern of ill health and disease in populations.⁹ Public health policy is largely founded upon such analyses, particularly where differences are inequitable in the sense of unjust. The central epidemiological question is this: why is a disease more common in one racial or ethnic group of people than another? For example, why in comparison with the British population as a whole is diabetes so common in people who originate in India but live in Britain, and yet why is colorectal cancer relatively uncommon? Answers to these questions would help understand better the causes of disease, and bring benefit to all populations. The mysteries behind the myriad of ethnic differences are, however, not easily solved. An emphasis on disease differences, so appropriate to the analysis required in the science of epidemiology, is deeply influential in the health policy and management arena where it is sometimes inappropriate for similarities may matter more.¹⁰ As interest in, and the influence of, research on ethnicity and race is rising it is important that the conceptual basis of the work is sound. The forces that will stimulate more work on ethnicity, race, and health include the new genetics, the focus on health and health care inequalities, globalisation, migration, and the increase in the movement of refugees and asylum seekers.

A PROBLEM WITH BASICS: CONCEPTS AND TERMINOLOGY RELATING TO ETHNICITY AND RACE

Ethnicity and race

As implied above the concepts of ethnicity and race are being hotly debated in epidemiology. Furthermore, there is no consensus on appropriate terms for use in the scientific study of health by ethnicity and race, and published guidelines on how to use these concepts, from a number of journals, are yet to be widely adopted.

The word ethnicity derives from the Greek word *ethnos*, meaning a nation. Ethnicity is a multi-faceted quality that refers to the group to which people belong, and/or are perceived to belong, as a result of certain shared characteristics, including geographical and ancestral origins, but particularly cultural traditions and languages. The characteristics that define ethnicity are not fixed or easily measured, so ethnicity

Correspondence to:
Professor R Bhopal,
Section of Public Health
Sciences, Medical School,
Teviot Place, Edinburgh
EH8 9AG, UK;
Raj.Bhopal@ed.ac.uk

Accepted for publication
18 December 2003

is imprecise and fluid. Ethnicity differs from race, nationality, religion, and migrant status, sometimes in subtle ways, but may include facets of these other concepts. It follows that investigators who wish to study ethnicity should collect data on such underlying factors, especially language, religion, country of birth, and family origins.

The measurement, or assignment, of ethnicity is problematic despite much research and debate. Presently, self definition of ethnicity is gaining favour. The problem is that the self assessment changes over time and with context, though this fluidity also has strengths. The currently favoured groupings in the UK, for example, include white, Indian, and Pakistani. Such groupings hide massive within group heterogeneity, which diminishes the value of ethnic categorisation as a means of delivering culturally appropriate health care, and in understanding the causes of ethnic variations in disease. Such broad categories may not fit with self identity of ethnic group.

The biological concept of race, whereby human populations were divided into sub-species mainly on the basis of visible physical characteristics, was dominant from the early 19th century to its decline with the defeat of the Nazis at the end of the second world war.⁹ None of the numerous racial classifications have stood the test of time, though there are echoes in current classifications. In retrospect, the biological concept of race was ill defined, poorly understood, and invalid and the science based on it needed sharper scientific criticism.¹¹ The modern concept of race, particularly as utilised by many scholars in the United States, emphasises its social origins rather than its biological basis.^{12, 13} In this perspective, race provides a way of defining, for social purposes, populations that look different and have different ancestral roots. It is evident, however, that even this social concept of race is ultimately based on physical and hence biological factors, though these are de-emphasised in social epidemiology as secondary, unimportant matters. The term race should be used with caution for its history is one of misuse and injustice. In the study of racism, however, the concept is central.

While race and ethnicity are different, they are overlapping concepts that are often used synonymously, a trend fostered by the increasing use, particularly in the United States, of the compound word race/ethnicity. This trend is unfortunate, but purposefully reflects the conceptual and practical problems of separating the two concepts, and traditions of race based analysis in the USA. In contrast, in Europe race is being abandoned in favour of ethnicity. Worldwide ethnicity is in the ascendancy.¹⁴ Whatever terms authors use, the underlying concepts ought to be discussed. In practice, a clear definition of what is meant by the terms ethnicity and race in publications is often lacking making it difficult to compare studies, particularly internationally. This is not surprising as race and ethnicity are complex, multidimensional concepts changing with time and therefore subject to varying interpretations. The field has, however, been weakened by inconsistent use of terms and concepts.¹⁵

Categories and classifications

To put the concepts of race and ethnicity into operation we need to derive categories, as most clearly seen in national censuses. Researchers have mostly used such administrative categories for race and ethnicity, even when these are acknowledged by those developing them as having no scientific or anthropological validity.¹⁶ This unsatisfactory state of affairs can only be remedied if scientists become more deeply involved in the development of categories, and are not merely end users. Scientists' use of existing classifications can be interpreted as an endorsement of their validity. As a minimum, researchers should explain their

understanding of the concepts of race or ethnicity and the classification they use (for example, one or a mix of ancestry, geographical origin, birthplace, language, religion, migration history, name, self identity, observation, etc).

It is easy to forget that categories are merely labels, and no more than a first step to understanding and defining a person's ethnicity or race. Such labels need to be recognised as shorthand for potentially important information. Authors should be describing the characteristics of the populations they are referring to. For example, the label "South Asian" should not be used if the population referred to is Bangladeshi. Bangladeshis are different from other South Asian populations. For example, Bangladeshi men have an extremely high prevalence of smoking, while some other South Asian groups of men (particularly Sikhs, but also Hindus) have a low prevalence, a vital fact lost by studies of smoking in "South Asians" combined.¹⁷ Popular terminology for ethnic minority populations (Asians, Blacks, Chinese, etc) may suffice for everyday conversation or political exchange but is too crude for research, and when used needs accurate definition.^{1, 15}

Editors are responsible for ensuring scientific rigour and high quality writing in their journals, yet few journals or books have appropriate policies that are implemented vigorously. This article may spur the *JECH* to adopt a policy on this matter. Achieving conceptual and terminological accuracy remains an important and challenging goal for authors and editors alike.

Ethnicity and health research requires a consideration of concepts and terms beyond those referring to race and ethnicity, for example, culture, acculturation, migration, asylum, refugee, equity etc—but these are beyond this paper and some are covered in other glossaries in *JECH*.

With this background, I offer a glossary as an example, but it needs development in terms of geographical specificity, scope, and precision. The terms and definitions are an amalgam of the concepts of race and ethnicity. This glossary is an enlarged version of one I have used for some years with some success. Clearly, using the principles evident in the glossary, for example, giving primacy to self identity while acknowledging related factors such as ancestry and the link to racial classifications, additional terms for other ethnic and racial groups internationally can be added. It may be that contributions from *JECH* readers will lead to a deeper and longer glossary with truly international applications.

EXAMPLE OF A GLOSSARY OF TERMS RELATING TO ETHNICITY AND RACE

Asian

Strictly, this label applies to anyone originating from the Asian continent. In practice, this term is used in the United Kingdom to mean people with ancestry in the Indian subcontinent. In the United States, the term has broader meaning, but is mostly used to denote people of far Eastern origins, for example, Chinese, Japanese, and Filipinos. More specific terms should be used whenever possible.

Asian Indian

A term currently used synonymously with South Asian (see below), but with the important limitation that major South Asian populations such as Pakistani and Bangladeshi may not identify with it. This term is being used in North America to distinguish the population from Native Americans, previously known as American Indians.

African

A person with African ancestral origins who self identifies, or is identified, as African, but excluding those of other ancestry, for example, European and South Asian. This term

is the currently preferred description for more specific categories, as in African American, for example. (In terms of racial classifications, this population approximates to the group historically known as Negroid or similar terms.) In practice, Northern Africans from Algeria, Morocco, and such countries are excluded from this category. (See also Black.)

Afro-Caribbean/African Caribbean

A person of African ancestral origins whose family settled in the Caribbean before emigrating and who self identifies, or is identified, as Afro-Caribbean (in terms of racial classifications, this population approximates to the group known as Negroid or similar terms). (See also Black.)

Bangladeshi

A person whose ancestry lies in the Indian subcontinent who self identifies, or is identified, as Bangladeshi. (See also South Asian.) Between 1947 and 1971 the land known as Bangladesh was East Pakistan and before that India. There is no clear cut equivalent in terms of racial classifications, though historically Northern Indians have been classified as white, and some Indian tribes as aboriginal. (The racial term Malayan, coined by Blumenbach, is forgotten as purposeless.)

Black

A person with African ancestral origins, who self identifies, or is identified, as Black, African or Afro-Caribbean (see, African and Afro-Caribbean). The word is capitalised to signify its specific use in this way. In some circumstances the word Black signifies all non-white minority populations, and in this use serves political purposes. While this term was widely supported in the late 20th century there are signs that such support is diminishing.

Caucasian

An Indo-European. This is Blumenbach's 18th century term for the white race of mankind, which he derived from the people who lived in the Caucasus. This term is usually used synonymously with Caucasoid, European, or White. Alone among terms derived from traditional racial classification, Caucasian remains popular in both science and everyday language.

Chinese

A person with ancestral origins in China, who self identifies, or is identified, as Chinese. (In terms of historical racial classifications, Chinese approximate to the group known as Mongolian or Mongoloid.)

Ethnicity

The social group a person belongs to, and either identifies with or is identified with by others, as a result of a mix of cultural and other factors including language, diet, religion, ancestry, and physical features traditionally associated with race (see race). Increasingly, the concept is being used synonymously with race but the trend is pragmatic rather than scientific.

Ethnic minority group

Usually, but not always, this phrase is used to refer to a non-white population. Alternatively, it may be used to describe a specific identifiable group, for example, gypsy travellers, and less commonly, Irish in the UK. Some people consider the phrase inaccurate and prefer minority ethnic group, but the two phrases are used synonymously.

Ethnocentrism

The tendency to perceive and interpret from the standpoint of one's own culture. In epidemiology the tendency is reflected

in the practice of using the White population as the norm or standard (see White).

European

European primarily means an inhabitant of Europe, or one with ancestral origins in Europe. Effectively this is used in epidemiology and public health as a synonym for White (see below). Europeans are placed in the racial classification Caucasian, more recently known as Europid (the latter has not proved popular).

General population

Everyone in the population being studied, irrespective of race or ethnicity.

Hindu

An old, now seldom used term, for Indians. A term occasionally used more or less synonymously with South Asian. In some countries such as Holland the term is used to describe the ethnicity of Surinamese of Indian subcontinent ancestry.

Hispanic

A person of Latin American descent (with some degree of Spanish or Portuguese ancestral origins), who self identifies, or is identified, as Hispanic irrespective of other racial or ethnic considerations. In the United States this term, often used interchangeably with Latino, is considered an indicator of ethnic origin.

Indian

A person whose ancestry lies in the Indian sub-continent who identifies, or is identified, as Indian (see, South Asian). (Major changes to India's geographical boundaries took place in 1947 when Pakistan was created.)

Indigenous

This term is usually used to mean a person who belongs naturally to a place in the sense of long term family origins (see Native). This term is sometimes used to identify the majority population, for example, in the United Kingdom as an alternative to the word White. In some parts of the world, for example, Australia, the word indigenous is used specifically to refer to aboriginal populations (for example, Aborigine).

Institutional racism (See racism)

Irish

A person whose ancestry lies in Ireland who self identifies or is identified, as Irish but this label is generally restricted to the White population (see, White).

Majority population

When used in race/ethnicity studies this phrase is usually used as a synonym for White or European.

Minority ethnic group

See ethnic minority group. Increasingly used as the preferred phrase and replacing ethnic minority group.

Native

Sometimes this word is used to refer to populations born, or with family origins, in a place (see indigenous). This was also a pejorative term meaning populations belonging to a non-European and imperfectly civilised or savage race, so writers need to take care.

Non-Asian/Non-Chinese, etc

This type of term is rarely defined but self evidently implies those not belonging to the group under study. This degree of non-specificity is not recommended.

Occidental

This is a very rarely used term meaning a native or inhabitant of the Occident (West), and effectively a synonym for European, but readers need to be aware of it as the antonym of Oriental.

Oriental

A term meaning a native or inhabitant of the Orient (East). This term is in occasional use in epidemiology, usually referring to Far Eastern populations. It is too general to be useful.

Pakistani

A person whose ancestry lies in the Indian subcontinent who identifies, or is identified, as Pakistani (see South Asian). Some Pakistanis may have birth or ancestral roots in the current territory of India but identify with Pakistan, a country created in 1947.

Race

By historical and common usage the group (sub-species in traditional scientific use) a person belongs to as a result of a mix of physical features such as skin colour and hair texture, which reflect ancestry and geographical origins, as identified by others or, increasingly, as self identified. The importance of social factors in the creation and perpetuation of racial categories has led to the concept broadening to include a common social and political heritage, making its use similar to ethnicity. Race and ethnicity are increasingly used as synonyms causing some confusion and leading to the hybrid terms race/ethnicity (see Ethnicity).

Racism/institutional racism

A belief that some races are superior to others, used to devise and justify individual and collective actions that create and sustain inequality among racial and ethnic groups. Individual racism is usually manifested in decisions and behaviours that disadvantage small numbers of people. Institutional racism, whereby policies and traditions, sometimes unwittingly, favour a particular racial or ethnic group, may be less obvious but may disadvantage large populations.

Racial prejudice

Negative beliefs, perceptions, or attitudes towards one or more ethnic or racial groups.

Reference/control/comparison

This refers to the standard against which a population that is being studied can be compared with to permit an analysis of similarities and differences. The concept is fundamental to epidemiology, and this terminology is preferable to non-specific ethnic or racial terms such as non-Asian, or general or even White population.

South Asian

A person whose ancestry is in the countries of the Indian subcontinent, including India, Pakistan, Bangladesh, and Sri Lanka (in terms of racial classifications, most people in this group probably fit best into Caucasian or Caucasoid but this is confusing and is not recommended). This label is usually assigned, for individuals rarely identify with it. (See also Indian, Indian Asian, Asian, Pakistani, Bangladeshi.)

Western

A person or populations with ancestry in a region conventionally known as the west, effectively European countries, as distinguished from Eastern or Oriental populations.

White

The term usually used to describe people with European ancestral origins who identify, or are identified, as White (sometimes called European, or in terms of racial classifications, the group known as Caucasian or Caucasoid). The

Some challenges for epidemiological research on ethnicity, race, and health

- Inclusion of minorities in research
- Clarification of the purpose of the research
- Definitions of concepts relating to ethnicity and race
- Definition and precision of terms, and ethnic/racial classifications
- Recognition of heterogeneity within ethnic minority groups
- Identification of representative populations
- Ensuring comparability of populations that are to be compared, requiring socioeconomic data over the life course
- Accurate measurement of the denominators and numerators, in calculating rates
- Ensuring the quality of data, particularly in cross cultural comparability
- Maximising completeness of data collection
- Avoiding misinterpretation of differences that are attributable to confounding variables
- Pinpointing genetic bases of genetic hypotheses
- Proper interpretations of associations as causal or non-causal
- Maximising validity and generalisability of the research
- Presentation of research to achieve benefits for the population studied, and avoid stigmatisation and racism
- Appropriate action to follow the research

word is capitalised to highlight its specific use. The term has served to distinguish these groups from those groups with skin of other colours (black, yellow, etc), and hence derives from the concept of race but is used as an indicator of ethnicity. There are problems of poverty and excess disease in subgroups of the White population, which cannot be unearthed and tackled by using the label White.

Mixed and other race or ethnic group

This glossary omits a clear exposition on these terms, which require fresh thought. The increasing importance of the category mixed (ethnicity or race) is self evident. The increasing acceptance of sexual unions that cross ethnic and racial boundaries is adding both richness and complexity to most societies. The way to categorise people born of such unions is unclear and the current approaches are inadequate, partly because the number of potential categories is huge. Another category seen in racial classifications is "other", this permitting those not included to identify themselves, or be identified by the observer. In both instances the solution is, most probably, to offer space for free text responses for people to identify themselves. These responses, however, need to be coded, analysed, summarised, quantified, and published. Without this individually small, but collectively large, populations remain hidden when policy on ethnic diversity is made.

CONCLUSIONS

As this paper shows even basic work such as defining terms is problematic, and the challenges are compounded by the pace of social change, and scientific practice.

Ethnicity is replacing the scientifically limited and somewhat discredited term race in the scientific literature but the latter concept is necessary to study racism.¹⁴ The focus of work on race and ethnicity tends to be on those populations with comparatively adverse health outcomes. Clearly, it is not only non-white minority ethnic groups that are in this position. There is a case for conceptualising, categorising, and studying more the White racial and ethnic sub-groups.¹⁸ In a future glossary there may be far more emphasis on White ethnic groups, a trend to be welcomed.

This paper has focused on the concepts and terminology used in the study of the health of minority ethnic and racial groups. It has introduced some of the fundamental debates. There are other issues, for example, whether international understanding and agreement on these concepts and terms, is achievable, the comparative health of population sub-groups within the populations defined by current categories, and empirical demonstration that the benefits of data by ethnicity and race exceed the costs, in particular that they help improve the health status of the study populations. The box summarises some of the challenges facing epidemiology in ethnicity and health.

Purpose and context are the prime determinants of the way that race and ethnicity concepts are applied, classifications are devised and used, and data are analysed and presented.¹⁹ This paper has tackled this topic mainly from the perspective of epidemiology and public health, and from a UK context, though infused by ideas from North America and some Northern European countries. The same challenge needs to be tackled from the perspective of other purposes and other nations. Then we would have a foundation for an internationally acceptable and more comprehensive glossary. The task is enormous but if the subject is to mature it needs to be tackled. In the field of health there is a case for leadership from a partnership including the WHO, International Epidemiological Association, and an organisation such as the World Association of Medical Editors.

ACKNOWLEDGEMENTS

I thank the editors of *JECH* for commissioning this paper, and four anonymous referees for their detailed, thoughtful, and critical feedback, which led to a substantial revision. This paper builds on my own and others' work, much of it listed in the bibliography. I acknowledge these debts. Professor Aziz Sheikh gave useful feedback on an earlier draft.



Please visit the journal web site (<http://www.jech.com/supplemental>) for a fuller bibliography on this subject.

APPENDIX

An example of an introductory note on ethnicity used by the author in past publications, which may be adapted by other authors.

A NOTE ON TERMINOLOGY RELATING TO ETHNICITY

There is no consensus on appropriate terms for the scientific study of health by ethnicity, and published guidelines are yet to be widely adopted. We have followed general conventions used in the UK and, whenever appropriate, the terminology used by the original authors. For example, in the UK the term ethnic minority group usually refers to minority populations of non-European origin and characterised by their non-white status. (We use it this way here). The term South Asian refers to populations originating from the Indian Sub-continent, effectively, India, Pakistan, Bangladesh, and Sri Lanka. White is the term currently used to describe people with European ancestral origins. By ethnicity we mean the group a person belongs to as a result of a mix of cultural factors including language, diet, religion, and ancestry.

REFERENCES

- 1 **Bhopal RS**, Phillimore P, Kohli HS. Inappropriate use of the term 'Asian': an obstacle to ethnicity and health research. *J Public Health Med* 1991;**13**:244–6.
- 2 **Bhopal RS**. Future research on the health of ethnic minorities; back to basics. In: Ahmad WIU, ed. *The politics of 'race' and health*. Bradford: University of Bradford, 1992. [Re-printed].
- 3 **Bhopal RS**. Ethnicity. In: Boyd KM, et al, eds. *The new dictionary of medical ethics*. London: BMJ Books 1997:89–90.
- 4 **Ahdiieh L**, Hahn RA. Use of the terms 'race', 'ethnicity', and 'national origins': a review of articles in the *American Journal of Public Health*, 1980–1989. *Ethnicity and Health* 1996;**1**:95–8.
- 5 **Chaturvedi N**. Ethnicity as an epidemiological determinant—crudely racist or crucially important? *Int J Epidemiol* 2001;**30**:925–7.
- 6 **Cooper R**. A note on the biological concept of race and its application in epidemiological research. *Am Heart J* 1984;**108**:715–23.
- 7 **Hahn RA**, Stroup DF. Race and ethnicity in public health surveillance: criteria for the scientific use of social categories. *Public Health Rep* 1994;**109**:4–12.
- 8 **Osborne NG**, Feit MD. The use of race in medical research. *JAMA* 1992;**267**:275–9.
- 9 **Senior P**, Bhopal RS. Ethnicity as a variable in epidemiological research. *BMJ* 1994;**309**:327–9.
- 10 **Gill PS**, Kai J, Bhopal RS, et al. Health care needs assessment: black and minority ethnic groups. In: Raftery J, ed. *Health care needs assessment. The epidemiologically based needs assessment reviews*. 3rd series. Abingdon: Radcliffe Medical Press (in press). (Available in full at <http://hcna.radcliffe-oxford.com/bemgframe.htm>).
- 11 **Bhopal RS**. Is research into ethnicity and health racist, unsound, or important science? *BMJ* 1997;**314**:1751–6.
- 12 **Kaplan JB**, Bennett T. Use of race and ethnicity in biomedical publication. *JAMA* 2003;**289**:2709–16.
- 13 **President's Cancer Panel, Report of the Chairman**. *The meaning of race in science—considerations for cancer research*. Bethesda: National Institutes of Health, National Cancer Institute, 1997.
- 14 **Afshari R**, Bhopal RS. Changing pattern of use of 'ethnicity' and 'race' in scientific literature. *Int J Epidemiol* 2002;**31**:1074–76.
- 15 **Aspinall PJ**. Collective terminology to describe the minority ethnic population: the persistence of confusion and ambiguity in usage. *Sociology* 2002;**36**:803–16.
- 16 **Office of Management and Budget**. *Standards for the classification of Federal data on race and ethnicity*. Washington: Office of Management and Budget, 1994. (http://www.whitehouse.gov/omb/fedreg/notice_15.html).
- 17 **Bhopal RS**, Unwin N, White M, et al. Heterogeneity of coronary heart disease risk factors in Indian, Pakistani, Bangladeshi and European origin populations: cross sectional study. *BMJ* 1999;**319**:215–20.
- 18 **Bhopal R**, Donaldson L. White, European, Western, Caucasian or What? Inappropriate labelling in research on race, ethnicity and health. *Am J Public Health* 1998;**88**:1303–7.
- 19 **Bhopal RS**. Race and ethnicity as epidemiological variables: centrality of purpose and context. In: Macbeth H, ed. *Ethnicity and health*. London: Taylor and Francis, 2001:21–40.