

Assignment of ethnicity in patients with acute ischemic stroke in northern israel

Gregory Telman^{1*}, Efim Kouperberg², Moshe Herskovitz², Tarek Diab², Hilel Hurani², Elliot Sprecher²

¹Department of Neurology, Rambam Medical Center, Haifa, Israel; *Corresponding Author: g_telman@rambam.health.gov.il

²Department of Neurology, Technion Faculty of Medicine, Rambam Health Care Campus, Haifa, Israel.

Received 23 January 2010; revised 1 February 2010; accepted 11 March 2010.

ABSTRACT

The description of various methods for ethnicity classification can be found in the literature, though their reliability still remains unclear. We examined inter-observer agreement in defining the ethnic identification of patients in a bi-ethnic population (Arab-Jewish) in northern Israel, using place of birth and residence in addition to given and family names. Data about 1006 consecutive patients with acute ischemic stroke were gathered from our stroke registry. The data were analyzed by four independent observers (authors MH, TD, HH, GT) aiming to assign patients either as Arabs or Jews. Agreement between all four observers was excellent, as assessed by Fleiss' Kappa statistic ($\kappa = 0.96$). We conclude that the use of given and family names of patients, together with their place of birth and residence, achieved near-perfect inter-observer agreement and a highly reliable assignment of ethnicity in two large ethnic population groups – Arabs and Jews – in northern Israel.

Keywords: Stroke; Ethnicity

1. INTRODUCTION

Ethnic differences in the etiology, clinical picture, work-up, and outcome of numerous diseases are well established [1-9]. This problem seems to be attracting growing interest. One of the important issues in studying ethnic influence is the correct assignment of the ethnic identity of patients involved in any given study. Obviously, precise ethnic identification is required for obtaining reliable results. The description of various methods for ethnicity classification can be found in the literature, though their reliability still remains unclear [10-15]. Therefore, we conducted an inter-observer agreement study on the ethnic assignment of patients in a bi-ethnic population (Arab-Jewish) in northern Israel,

using place of birth and residence in addition to given and family names. The results of this study are highly important for planning of future studies devoted to the influence of ethnicity on different aspects of cerebrovascular diseases in our region.

2. MATERIALS AND METHODS

Data were gathered from our stroke registry on 1006 consecutive patients with acute stroke who had been hospitalized once in the Department of Neurology in a major hospital in northern Israel (Rambam Health Care Campus) during the period from 2001 to 2003. For the ease of processing and the prevention of double-counts, 25 additional patients who had two or more hospitalizations during this period were not included in the study. The data presented for analysis included the patients' given and family names together with their place of birth and residence at the time of the event.

Four independent observers, all of whom are medical doctors, were invited to participate in the analysis and to define an ethnic assignment of patients as Arab or Jewish. The observers included an Israeli-born Jew (M.H), two Israeli-born Arabs (T.D and H.H), and one immigrant from the former Soviet Union living in Israel since 1992 (G.T). All of the observers received a table of data prepared by an additional person (E.K), who was not serving as an observer. The table included four items: first and family names together with place of birth and residence at the time of the index event. The observers were required to assign each patient as an Arab or a Jew. After the assignments were completed by all four observers, the statistical analysis was conducted. Since four observers were employed, Fleiss' Kappa agreement statistic was used instead of the conventional two-observer Kappa.

3. RESULTS

Results of estimation of ethnicity by every observer are presented in **Table 1**. Note that complete agreement

Table 1. Distribution of agreement on ethnic assignment across observers.

Observer A	Observer B	Observer C	Observer D	Number of cases
J	J	J	J	786
A	A	A	A	195
J	A	J	J	6
A	A	A	J	4
J	J	J	A	4
A	A	J	A	3
A	J	J	A	2
J	J	A	J	2
A	J	A	A	1
A	J	A	J	1
J	A	A	A	1
J	A	A	J	1

J = Jewish ethnicity; A = Arab ethnicity.

across four observers was found in more than 97.5% of the cases evaluated.

Fleiss' Kappa was 0.96 across all four observers, indicating near-perfect agreement in the assessment of ethnicity.

4. DISCUSSION

The identification of patient ethnicity by names, with or without additional factors such as place of birth, has been widely used in medicine for many years [16-21]. There have also been numerous attempts to use computerized systems for this aim as well [22-25]. However, these methods achieve results of varying reliability when applied in different ethnic populations [26-30].

We attempted to assign ethnicity in a bi-ethnic population of Arabs and Jews in northern Israel. Both of these large ethnic groups are ethnically heterogeneous. The Jewish population includes two main groups: Ashkenazi Jews from Central and Eastern Europe, and Sephardic Jews from the Mediterranean and North Africa [31]. The Arab population includes Sunni Muslims, Druze, and Christian Arabs [32].

Despite being so ethnically heterogeneous, the first names and surnames of both populations remain highly distinctive between them. Two additional circumstances are helpful in the assignment of ethnicity to Arabs and Jews in the Israeli population. The first circumstance is that a significant number of Jewish people immigrated to Israel before and especially after the State of Israel was established in 1948. Thus, those patients who were born

in countries other than Israel can safely be assigned to the Jewish population. The second circumstance is that there are numerous densely populated areas of Israeli Arabs. As there is very little genetic intermingling between the two populations, a patient's place of residence in one of these areas is a clear indicator of Arab ethnicity. In conclusion, when considering the patients' given and family names, together with their place of birth and residence, we were able to achieve extremely good inter-observer agreement, thereby allowing a highly reliable assignment of ethnicity in the Arab-Jewish population of northern Israel. The results of our study demonstrate that even in case of retrospective studies, when there is no possibility to define ethnicity by direct patient interview, the method of ethnicity assignment by names and place of birth could be highly reliable.

REFERENCES

- [1] Lee, Y.T. (1981) Cancer statistics of Chinese versus Americans. *Journal of Surgical Oncology*, **17**, 355-366. doi:10.1002/jso.2930170408
- [2] Rogers, R.G. (1989) Ethnic and birth weight differences in cause-specific infant mortality. *Demography*, **26**, 335-343. doi:10.2307/2061530
- [3] Parker, K.D., Calhoun, T. and Weaver, G. (2000) Variables associated with adolescent alcohol use: A multi-ethnic comparison. *The Journal of Social Psychology*, **140**, 51-62. doi:10.1080/00224540009600445
- [4] Rosner, B., Prineas, R., Daniels, S.R. and Loggie, J. (2000) Blood pressure differences between blacks and whites in relation to body size among US children and adolescents. *American Journal of Epidemiology*, **10**, 1007-1019.
- [5] Connell, P., McKeivitt, C. and Low, N. (2004) Investigating ethnic differences in sexual health: Focus groups with young people. *Sexually Transmitted Infections*, **80**, 300-305. doi:10.1136/sti.2003.005181
- [6] Tan, G., Jensen, M.P. and Thornby, J. (2005) Anderson KO. Ethnicity, control appraisal, coping, and adjustment to chronic pain among black and white Americans. *Pain Medicine*, **6**, 18-28. doi:10.1111/j.1526-4637.2005.05008.x
- [7] Rice, V.M. (2005) Strategies and issues for managing menopause-related symptoms in diverse populations: Ethnic and racial diversity. *American Journal of Medicine*, **118**, 142-147. doi:10.1016/j.amjmed.2005.09.048
- [8] Ikegami, H., Fujisawa, T., et al. (2006) Genetics of type 1 diabetes: Similarities and differences between Asian and Caucasian populations. *The New York Academy of Sciences*, **1079**, 51-59. doi:10.1196/annals.1375.008
- [9] Miller, M.A., et al. (2009) Ethnic and sex differences in circulating endotoxin levels: A novel marker of atherosclerotic and cardiovascular risk in a British multi-ethnic population. *Atherosclerosis*, **203**, 494-502. doi:10.1016/j.atherosclerosis.2008.06.018
- [10] Boydell, J., et al. (2001) Incidence of schizophrenia in ethnic minorities in London: ecological study into interactions with environment. *British Medical Journal*, **323**,

- 1336-1338. [doi:10.1136/bmj.323.7325.1336](https://doi.org/10.1136/bmj.323.7325.1336)
- [11] Hazuda, H.P., et al. (1986) A comparison of three indicators for identifying Mexican Americans in epidemiologic research. Methodological Findings from the San Antonio Heart Study. *American Journal of Epidemiology*, **123**, 96-112.
- [12] Stewart, S.L., et al. (1999) Comparison of methods for classifying Hispanic ethnicity in a population-based cancer registry. *American Journal of Epidemiology*, **149**, 1063-1071.
- [13] Senior, P.A. and Bhopal, R. (1994) Ethnicity as a variable in epidemiological research. *British Medical Journal*, **309**, 327-330.
- [14] Ford, M.E. and Kelly, P.A. (2005) Conceptualizing and categorizing race and ethnicity in health services research. *Health Services Research*, **40**, 1658-1675. [doi:10.1111/j.1475-6773.2005.00449.x](https://doi.org/10.1111/j.1475-6773.2005.00449.x)
- [15] Gomez, S.L. and Glaser, S.L. (2006) Misclassification of race/ethnicity in a population-based cancer registry (United States). *Cancer Causes Control*, **17**, 771-781. [doi:10.1007/s10552-006-0013-y](https://doi.org/10.1007/s10552-006-0013-y)
- [16] Nicoll, A., Bassett, K. and Ulijaszek, S.J. (1986) What's in a name? Accuracy of using surnames and forenames in ascribing Asian ethnic identity in English populations. *Journal of Epidemiology & Community Health*, **40**, 364-368. [doi:10.1136/jech.40.4.364](https://doi.org/10.1136/jech.40.4.364)
- [17] Coldman, A.J., Braun, T. and Gallagher, R.P. (1988) The classification of ethnic status using name information. *Journal of Epidemiology & Community Health*, **42**, 390-395. [doi:10.1136/jech.42.4.390](https://doi.org/10.1136/jech.42.4.390)
- [18] Gupta, M., et al. (2002) Risk factors, hospital management and outcomes after acute myocardial infarction in South Asian Canadians and matched control subjects. *Canadian Medical Association Journal*, **166**, 717-722.
- [19] Schwartz, K.L., et al. (2004) Cancer among Arab Americans in the metropolitan Detroit area. *Ethnicity & Disease*, **14**, 141-146.
- [20] Hensley Alford, S., et al. (2009) Breast cancer characteristics at diagnosis and survival among Arab-American women compared to European- and African-American women. *Breast Cancer Research and Treatment*, **114**, 339-346. [doi:10.1007/s10549-008-9999-z](https://doi.org/10.1007/s10549-008-9999-z)
- [21] Coronado, G.D., et al. (2002) Assessing cervical cancer risk in Hispanics. *Cancer Epidemiology, Biomarkers & Prevention*, **11**, 979-984.
- [22] Harding, S., Dews, H. and Simpson, S.L. (1999) The potential to identify South Asians using a computerised algorithm to classify names. *Most Popular Trends*, **97**, 46-49.
- [23] Macfarlane, G.J., et al. (2007) Determining aspects of ethnicity amongst persons of South Asian origin: The use of a surname- classification programme (Nam Pehchan). *Public Health*, **121**, 231-236. [doi:10.1016/j.puhe.2006.07.001](https://doi.org/10.1016/j.puhe.2006.07.001)
- [24] Nanchahal, K., et al. (2001) Development and validation of a computerized South Asian Names and Group Recognition Algorithm (SANGRA) for use in British health-related studies. *Journal of Public Health*, **23**, 278-285. [doi:10.1093/pubmed/23.4.278](https://doi.org/10.1093/pubmed/23.4.278)
- [25] Brant, L.J. and Boxall, E. (2009) The problem with using computer programmes to assign ethnicity: immigration decreases sensitivity. *Public Health*, **123**, 316-320. [doi:10.1016/j.puhe.2009.02.002](https://doi.org/10.1016/j.puhe.2009.02.002)
- [26] Nicoll, A., Bassett, K. and Ulijaszek, S.J. (1986) What's in a name? Accuracy of using surnames and forenames in ascribing Asian ethnic identity in English populations. *Journal of Epidemiology & Community Health*, **40**, 364-368. [doi:10.1136/jech.40.4.364](https://doi.org/10.1136/jech.40.4.364)
- [27] Swallen, K.C., et al. (1998) Accuracy of racial classification of Vietnamese patients in a population-based cancer registry. *Ethnicity & Disease*, **8**, 218-227.
- [28] Sheth, T., et al. (1997) Classifying ethnicity utilizing the Canadian Mortality Data Base. *Ethnicity & Health*, **2**, 287-295. [doi:10.1080/13557858.1997.9961837](https://doi.org/10.1080/13557858.1997.9961837)
- [29] Lagerberg, D., Magnusson, M. and Sundelin, C. (2005) Surname as a marker of ethnicity. A study from child health services shows that immigrant respective Swedish families seem to be isolated in different ways. *Lakartidningen*, **102**, 2145-2148.
- [30] Bouwhuis, C.B. and Moll, H.A. (2003) Determination of ethnicity in children in The Netherlands: Two methods compared. *European Journal of Epidemiology*, **18**, 385-358. [doi:10.1023/A:1024205226239](https://doi.org/10.1023/A:1024205226239)
- [31] Amar, A., et al. (1999) Molecular analysis of HLA class II polymorphisms among different ethnic groups in Israel. *Human Immunology*, **60**, 723-730. [doi:10.1016/S0198-8859\(99\)00043-9](https://doi.org/10.1016/S0198-8859(99)00043-9)
- [32] Rozen, P., et al. (2007) The changing incidence and sites of colorectal cancer in the Israeli Arab population and their clinical implications. *International Journal of Cancer*, **120**, 147-151. [doi:10.1002/ijc.22141](https://doi.org/10.1002/ijc.22141)