

30. Korn A: Coverage issues raised by comparisons between CPS and establishment employment. Proceedings of the Social Statistics Section, Part I, American Statistical Association, 1978, pp 60-69.
31. Reyes J, *et al*: The survey design for a residential survey of illegal aliens. Report submitted to the Immigration and Naturalization Services, Washington, DC September 1977.
32. Warren R: Immigration to the US. Hearing before the select committee on population, US House of Representatives, 95th Congress, Second Session, April 4-7, 1978, No. 5, pp 704-706.
33. Lancaster C, Scheuren FJ: Counting the uncountable illegals: some initial statistical speculations employing capture-recapture techniques. Proceedings of the Social Statistics Section, 1977: Part I American Statistical Association, 1978, pp 530-535.
34. Lesko Associates: Final Report: Basic Data and Guidance Required to implement a Major Illegal Alien Study during Fiscal Year 1976. Prepared for the Office of Planning and Evaluation. US Immigration and Naturalization Services, Washington, DC, 1975.
35. Robinson GJ: Estimating the approximate size of the illegal alien population in the United States by the comparative trend analysis of age-specific death rates. Paper presented at the annual meeting of the Population Association of America, Philadelphia, PA, April 26-28, 1979.
36. US Bureau of the Census: *Factfinder for the Nation*, CFF No. 1 (Rev.), Minority Statistics, June 1979.
37. Powell-Griner E, Streck D: A closer examination of neonatal mortality rates among the Texas Spanish surname population. *Am J Public Health* 1982; 72:993-999.
38. Trevino FM: Vital and health statistics for the US Hispanic population (editorial). *Am J Public Health* 1982; 72:979-982.

## On Comparing Studies of Different Raza Populations

DAVID E. HAYES-BAUTISTA, PhD

### *The Problem of Comparability of Studies*

The previous paper by Giachello, *et al*,<sup>1</sup> discusses the fact that, at various times, the Census has employed different definitions for the somewhat elusive "Hispanic" population, and that problems might be encountered in comparing older studies, using one definition, with more recent studies, using different definitions. The problem becomes more complex if, as Giachello suggests, one combines census data with other data sources.

As research monies become more scarce, there is the likely possibility that planners and administrators who wish to develop or modify programs to handle the increased "Raza"<sup>2</sup> patient load will not be able to mount local research efforts. Increasingly, such planners, administrators, and policy formulators will have to rely upon studies done elsewhere, at other times, to serve as a basis for health services planning.

As researchers are discovering, the Raza population is very heterogeneous, ranging from the non-Spanish-speaking Indians of rural Mexico to British descended Argentines or Germanic Chileans. There are racial, cultural, and language

differences between the various Raza groups: Mexican, Puerto Rican, Cuban, Salvadoreno, Chileno, Brasileno, etc.

If one is unable to mount a local population research effort, and turns instead to studies already done on other segments of the Raza population in other parts of the country, the problem is one of assessing the comparability and applicability of such studies to a local population of "Hispanics". If all a planner has available is research done on urban Cuban residents in Florida, can he/she apply the results to a rural, Mexican migrant population in Colorado?

A typology is offered here to assist in assessing the comparability of studies of the different Raza population groups to one another.

### *Operationalization of Definitions*

There are many different terms used to denote Raza populations (e.g., Spanish-American, Latino, Hispanic, Mexican-American, Puerto Rican, Spanish-speaking, etc.). There is no uniformity of terms and, to complicate matters, each may be operationalized differently. There may be one study of "Mexican-Americans" whose population consists of those who speak Spanish, and a different study of "Hispanics" whose population was self-identified.

While there are many specific ways of operationalizing different terms for Raza, they may be categorized into four types:

- Nationality, i.e., the country of origin of self or parents;
- Race (not used recently);

Address reprint requests to David E. Hayes-Bautista, PhD, Associate Professor, Department of Social and Administrative Health Sciences, School of Public Health, University of California, Berkeley, CA 94720. This paper, submitted to the *Journal* May 12, 1982, was revised and accepted for publication July 1, 1982.

**Editor's Note:** See also related article p 266 this issue.

Population Denominator		Income	Education	Occupation	Immigration	Rural/Urban
1. CULTURE						
a. Language Spanish-Speaking						
b. Spanish Surname	CASE I Rural California	low	low	farm worker	recent possibly undocumented	rural
2. NATIONALITY	CASE II Urban Florida Cuban	high	high	white collar	1st, 2nd, 3rd generation all legal	urban
3. RACE	CASE III Rural Colorado					
4. SELF-DEFINED		low	low	farm worker	recent possibly undocumented	rural

FIGURE 1—A Typology for Determining the Comparability of Studies of Different Raza Populations

- Culture, i.e., surname, language used, etc.;
- Phenomenological definitions, i.e., self-definitions.

Each one of these types of operationalization has its uses and limitations.

#### Nationality

This is perhaps the easiest trait to identify, as it takes a relatively simple answer to a simple question. One merely asks the birthplace of self, parents, or grandparents.

There are, however, many limitations: beyond the birthplace of self, intermarriage may complicate the way responses have to be handled. Furthermore, the question assumes a homogeneous national population composition, although some countries are so ethnically diverse that they are officially bi-lingual (e.g., Peru is officially a Quechua-Spanish-speaking country). Thus, many of the finer points for population studies may be lost.

#### Race

It is probably safe to assume that the majority of Raza populations have a significant Native American genetic background. It is not currently known if this group constitutes a single "race".

In many national and local studies from the 1930s to the 1950s, "Mexican" was considered a separate race. Gradually, the definition changed to make Mexicans "White persons of Spanish surname." The issue of race is almost always skirted in studies of Raza populations, but the bald fact is that Latinos are racially quite heterogeneous upon an Indian substrate ranging from pure Indian to pure White to pure Black, to pure Asian, with nearly every conceivable combination of *mestizo*.

One might wish to trace race for health research reasons, for example to look for genetically linked illness. One would then have to develop a way of separating race from

either culture or nationality to produce such "clean" samples.

While, in the past, mere observable characteristics were thought sufficient to classify a person racially as "Mexican", for current health services research, some form of genotyping may be indicated. Expense makes this option not viable for survey research. However, if one controls for nationality and income, one can develop a population that has a higher proportion of genetically similar individuals. Some countries (e.g., Mexico, Guatemala, Bolivia, Peru) have large proportions of Native Americans in the populations, while the population of other countries (e.g., Argentina, Chile, Uruguay) have almost entirely European backgrounds. Usually, income is highly correlated with race—Indians at the lower end of the scale, Europeans at the top. Thus, one could construct a rough racial sample knowing nationality, income, education, or some other indicator.

#### Cultural

Since the 1960s, cultural euphemisms have been utilized. These variables are subject to rapid change, due to acculturation and assimilation. Thus, the physical composition (nationality and race) of a research population may change rapidly over time, with the same operationalization being used. For example, the Spanish surname, Spanish language population in San Francisco in 1950 was nearly exclusively Mexican, but in 1980 is largely Central and South American.<sup>3</sup>

#### Self-Identification

While other operationalizations depend upon possession of some verifiable trait, the self-identification is more openly phenomenological, i.e., it relies upon the individual's perception of self. While there is some evidence that this

type of operationalization captures fewer people,<sup>4</sup> there is no definitive proof of this.

The instrument in Figure 1 provides a typology for assessing the comparability of studies using different operationalizations of Raza populations. It is a heuristic device to allow a planner to lay out systematically the characteristics of a given Raza population, and to compare them with those of another. It is intended to help a planner realize the extent of the heterogeneity of different Raza populations, and to help take this heterogeneity into account before assuming comparability of different samples.

To illustrate the use of Figure 1, let us follow the example of a health planner who needs health service utilization data for a Raza population in Colorado. Two utilization studies are available: one done on Spanish-surnamed farm workers in California and one on urban Cuban residents in Miami. The first step is to lay out the elements to be compared. The population denominators of the two studies are identified, Spanish surname in the rural California study (Case I) and nationality in the Cuban Study (Case II). The relevant demographic variables from each study are placed in the appropriate cell (income, education, occupation, immigration status, and rural/urban location).

The second step is to become sensitized to the local Raza population in order to gain some rough idea of how it identifies itself and what its demographic variables look like. An inventory of merchants, local newspapers, and conversations with key informants (particularly those in community-based organizations) should give a rough idea of how the population sees itself. These qualitative data may then be supplemented with quantitative data developed by local community-based organizations, or local community or state college research (e.g., student papers, etc.). The planner can then develop a "master" for the local area, specifying the major horizontal characteristics. Figure 1 is thus filled out, for a self-defined Mexican, low-income, low-education farm worker, rural population of recent, dubious immigration status (Case III).

The applicability of the other studies may then be roughly gauged, by taking into consideration the following decision rules:

If all the cells match, the applicability is likely to be quite high. In the cases given, the rural California and the rural Colorado population characteristics are practically

identical. Differences in utilization are more likely to be found in local structural characteristics (e.g., quantity and quality of services) than in population difference.

If none of the cells match, the application probably should not be attempted. The Cuban sample differs greatly: it is high income, high education, white collar, urban, of long-standing and legal immigration status.

If the cells match in part, but not in others, the critical ones are most likely to be the education and income variables. The applicability will be questionable, and at times suspect. However, planners have to come up with plans, so the applicability should be explained in detail, making clear all assumptions of applicability, differences in the populations, and some rough "confidence factor" developed, even if it is only anecdotal and narrative. If the income and education variables are not comparable, the application probably should not be attempted.

### Conclusions

The 1980s have been touted as the "Decade of the 'Hispanic' ", yet it is also the decade of sharply decreased resources for research. Health planning and policy making for this population will have to rely on studies with varying definitions and operationalizations of the population. This typology is offered as an attempt to assist in comparing and applying studies done on one Raza group to another. The heterogeneity of the Raza population must be taken into account, or plans and policies may miss the mark widely, and, in the current fiscal climate, few mistakes can be tolerated—there may not be money for second attempts.

### REFERENCES

1. Giachello AL, Bell R, Aday LA, Andersen RM: Uses of the 1980 census for Hispanic health services research. *Am J Public Health* 1982; 72:266-274.
2. Hayes-Bautista DE: Identifying 'Hispanic' populations: the influence of research methodology upon public policy. (editorial) *Am J Public Health* 1980; 70:353-356.
3. State of California, Health and Welfare Agency, Employment Development Department: *California Labor Market Issues: Hispanics*, September 1981.
4. Alameda-Contra Costa Health Systems Plan, 1978-1983. Oakland: Alameda-Contra Costa Health Systems Agency, 1978.