Coll. Antropol. **26** (2002) 1: 1–9
UDC 572:616-036.22:314.7
Original scientific paper

Formulation and Validation of a Composite Migration History Score for Epidemiological Research in Mexican American Women

B. M. Chakraborty

Division of Epidemiology and Biostatistics, Department of Environmental Health, University of Cincinnati, Cincinnati, USA

A B S T R A C T

This research developed a composite migration history score, using data on birth places of subjects, and their parents and grand parents, and the length of residence in US from 390 overweight Mexican American women. The derived migration history score is validated through a factor analysis of the nine dimensional migration history variables. The importance of this methodology of defining migration history score for epidemiological research is discussed in the context of studying the role of migration -related health behavior changes and their consequent impact on risks of untoward health outcomes in migrants.

Introduction

Mexican Americans are the fastest growing minority population of the United States[1]. As a group, they are also at greater risk of morbidity and mortality than the general United States population[2]. Of even more of concern is that they are experiencing a rise in the prevalence of risk factors for coronary heart disease[3]. Thus, factors associated with risks of morbidity and mortality in the Mexican Americans have been of significant interest in public health policy statements as well as research[4,5].

Studies in Mexican Americans have shown that healthy behavior (such as physical activity, moderate alcohol consumption, tobacco avoidance, weight control and regular sleeping habit) is correlated with improved physical health[6] and has a positive influence on survival in middle-aged as well as elderly men and

Received for publication April 29, 2002

women[7]. Consequently, unhealthy behaviors are major factors contributing to an increased risk of cardiovascular diseases in Mexican Americans[8,9]. In addition, there is a literature, largely separate from the one mentioned above, showing that health behaviors change in response to adaptation to the host culture among immigrants[10–13]. Some authors[14] attributed these migration-history-related changes of health behaviors to stresses accompanying the process of »acculturation« into a larger society. Thus, a question can be raised as to whether: 1) the high prevalence of chronic disease risk factors in Mexican Americans is directly associated with their history of immigration (i.e., this ethnic association is a consequence of migration-related stress) independently of their health behaviors, or 2) the differences in chronic disease risk factors by immigration history are mediated through differences in health practices among subjects with different history of immigration (as measured by a composite score of migration status).

Answers to such epidemiological research questions depend upon use of an objective and readily useable questionnaire that would collect reliable and valid migration history information. This research demonstrates that by implementing a simple demographic questionnaire, data collected on birth places of subjects and their immediate ancestors along with the length of residence and generation of migration can yield a composite migration history score that is useful for epidemiological research among the Mexican American women.

Materials and Methods

Study population

Migration history information used in this investigation was collected as a part of the baseline data from an ongoing study of the University of Texas, Houston, Health Science Center and Baylor College of Medicine, entitled »Unidos en Salud: Weight loss for Mexican Americans«. »Unidos en Salud« was initiated in 1993, at which time 390 Mexican-American women aged 18 to 65 years old were recruited from several communities in southern Texas along the US/Mexico border. Only non-diabetic, non-pregnant women, who had body mass index (BMI) from 25 to 40 kg/m^2 (i.e., were overweight), were enrolled in the study. Other details of the sampling design of the weight loss study are described in Poston et al.[15].

Migration history variables

Data used for the present analysis was collected at the time when the subjects were recruited for the weight-loss study. Therefore, the study design is cross -sectional. Trained bilingual research assistants administered a demographic questionnaire, validated for Mexican American populations, to each subject. Participants had the opportunity to be interviewed in either English or Spanish. Migration history data includes birthplace of subject, parents and grandparents (Mexico- vs. US-born), and length of time in the US. In addition to the number of years of residence in the US, this last item also classified subjects as immigrants of the first, second, and third-or -more generations. Thus, for each subject a nine dimensional migration history information was obtained. These questionnaires also contained data on language preference (e.g., language preferred during watching television, listening to radio, reading newspaper, etc.). However, data collected on these later items resulted in a number of methodological problems. For example, opportunities for exposure to alternative languages for these activities were not noted for the community, and furthermore, the responses to the language preference questions were also internally inconsistent. These lead to the

decision not to use the language preference dimension as a part of scoring the migration history status of subjects.

Statistical analyses

Only one of the nine dimensions of the migration history variables (namely, the length of residence in the US) was continuous, while all remaining eight components were categorical. Because of extreme skewness of this continuous variable, length of residence was categorized in 3 classes (< 5 yrs, 5–9 yrs, and ≥ 10 yrs.). Along with the frequency distributions of these nine categorical variables, nonparametric correlations (Spearman's ρ) were computed for all pairs of variables to check their interdependence with one another. The factor analysis program of SPSS 9.1 software package was used with the option of varimax rotation to extract the factors from the nine dimensional migration history variables. The summary result of this factor analysis was contrasted with an alternative definition of a composite migration history score for the purpose of validation.

Results

Frequency distributions

The univariate frequency distributions of the nine migration history variables are shown in Table 1 with a scoring system, following the methods of Cantero et al.[16] and Markides et al.[17]. As seen from these distributions, country of birth item was missing for several subjects, particularly for the grandparental generation (e.g., 66 missing data for father's father's country of birth). To examine if this introduced any systematic bias, the frequency distributions of Table 1 were recomputed for the complete records (n = 303), which were not significantly different from the ones shown. This suggests that missing items represent a random loss of information, and not a source of any systematic bias.

As seen from these distributions, majority of the women (86.9%) had been residing in the Starr County for 10 or more years, and over 69% of them are 3rd or more generation immigrants. About two -third or more of the subjects and their parents are born in the US, and only a little over one-half of the grand parents are born outside the US (mostly in Mexico).

Definition of the composite score of migration history

Markides et al.[17] and Cantero et al.[16] used the summed score of components of migration information as the composite score of migration history. Their rationale was high internal reliability (approximately 96%). A simple summed score is statistically unjustifiable in the present analysis, since two of the components (years of residence and immigration status) have three categories, while the remaining six have two categories for each. Thus, a simple sum would have placed more emphasis on the first two. Furthermore, for 87 of the 390 subjects, data on at least one item was missing. The following method is therefore better suited for defining a composite migration history score. First, each component is standardized with mean zero and unit (1.0) standard deviation. For each individual, the average of the standardized scores from all available items defines the composite migration history score (MHS). This maximizes the number of subjects for whom the MHS variable can be defined, without introducing any systematic bias, as shown from the homogeneity of distributions in the full data and in the ones with complete information.

Figure 1 shows the cumulative distribution plot of this composite score of migration history. Also depicted in this figure are the tertiles of the distribution of this score that define the three categories

TABLE 1
MIGRATION HISTORY VARIABLES, SCORES ASSIGNED TO EACH CATEGORY AND THE FREQUENCY DISTRIBUTION OF SUBJECTS

Item	Category	Score	Frequency (%)
Years of residence	< 5 years	0	21 (5.4)
	5–9 years	1	30 (7.7)
	≥ 10 years	2	339 (86.9)
Immigration status	1st generation	0	109 (27.9)
	2nd generation	1	11 (2.8)
	≥ 3rd generation	2	270 (69.2)
Birthplace of subjects	Mexico/other	0	109 (27.9)
	USA	1	281 (72.1)
Father	Mexico/other	0	127 (33.6)
	USA	1	251 (66.4)
Mother	Mexico/other	0	145 (37.8)
	USA	1	239 (62.2)
Father's father	Mexico/other	0	177 (53.8)
	USA	1	152 (46.2)
Father's mother	Mexico/other	0	180 (54.2)
	USA	1	152 (45.8)
Mother's father	Mexico/other	0	192 (55.8)
	USA	1	152 (44.2)
Mother's mother	Mexico/other	0	192 (54.9)
	USA	1	158 (45.1)

of migration history: recent (lowest tertile = score less than of equal to –0.17), intermediate (middle tertile = score between –0.17 and 0.60), and long-term (highest tertile = score above 0.60).

Validation of the composite migration history score

Two methods of validation were used to justify the use of the above composite score. First, nonparametric correlation coefficients (Spearman's ρ) were computed between all pairs of components of migration-related variables. Spearman's ρ was used, as eight of the nine items are discrete. Table 2 shows these nonparametric correlation coefficients among all pairs of the nine-item migration history variables.

All 36 pairwise correlation coefficients are significant at $p < 0.01$ level. However, the correlations of years of residence with the remaining eight variables are somewhat weaker (0.143 to 0.311). The correlation coefficients between the remaining eight variables are comparatively much stronger (ρ between 0.405 to 0.967). The high and nearly equal correlation between components supports the notion that probably each component is equally important in characterizing migration history.

These high correlations also prompted the extraction of the first principal factor from the nine migration-history related components. The first factor, obtained from the factor analysis (with varimax rotation), with an eigenvalue of 5.505 ex-

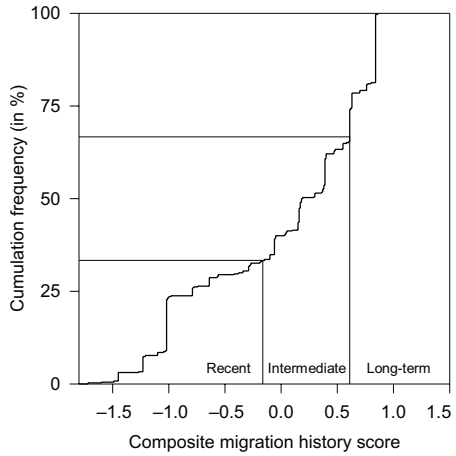

Fig. 1. Cumulative distribution function of the composite migration history score (average standardized scores of nine migration history variables), and definition of the three migration history categories of 390 Mexican American women by tertiles.

plained over 61% of the total variance. The next component accounted for only 11% of the variance. The scree plot suggested that one factor suffices the summarization of the nine item migration history variables. The coefficients of the component matrix for the first principal factor were:

- 0.257 for years of residence in Starr County;
- 0.932 for immigration status;
- 0.898 for subject's birthplace;
- 0.874 for father's birthplace;
- 0.838 for mother's birthplace;
- 0.755 for paternal grandfather's birthplace;
- 0.767 for paternal grandmother's birthplace;
- 0.746 for maternal grandfather's birthplace; and
- 0.764 for maternal grandmother's birthplace.

For each subject estimated factor scores were computed. Figure 2 shows a comparison of the estimated first principal factor score with the composite migration history score, obtained from the simple averages of the standardized component scores. This scatter plot ($r = 0.995$, $p <$

TABLE 2
NON-PARAMETRIC CORRELATIONS (SPEARMAN'S ρ) AMONG NINE MIGRATION HISTORY VARIABLES

	YRS	IMST	S-BP	F-BP	M-BP	FF-BP	FM-BP	MF-BP
IMST	0.292							
S-BP	0.311	0.967						
F-BP	0.248	0.855	0.796					
M-BP	0.212	0.766	0.713	0.645				
FF-BP	0.143	0.603	0.559	0.665	0.405			
FM-BP	0.188	0.602	0.568	0.653	0.474	0.737		
MF-BP	0.203	0.550	0.514	0.464	0.673	0.429	0.456	
MF-BP	0.158	0.585	0.548	0.485	0.694	0.423	0.511	0.755

All correlations are significant at 1% level.

YRS = years of residence in Starr County; S-BP = subject's birthplace; F-BP = father's birthplace; M-BP = mother's birthplace; FF-BP = father's father's birthplace; FM-BP = father's mother's birthplace; MF-BP = mother's father's birthplace; MM-BP = mother's mother's birthplace

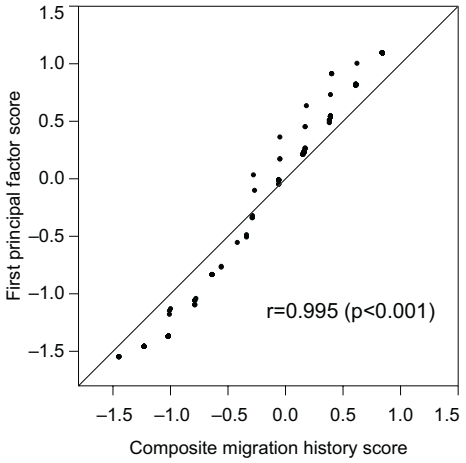

Fig. 2. Scatter plot of composite migration history score (average standardized scores of nine migration history variables) and the first principal factor of the nine variables.

0.001) suggests that the simple average of the standardized component scores is a valid migration history score, for quantitative as well as categorical analysis. In fact, tertiles defined based on the first principal factor classified the women exactly identically in recent, intermediate, and long-term immigrants, so that for all practical purposes the first principal factor and summed standardized scores are equivalent as the composite migration history score.

Discussion and Conclusion

Migration history of subjects has been used in epidemiological research for over four decades. Buell and Dunn[18] studied relative risks of mortality from cancer of the stomach, liver, and colon among Japanese men in Japan, Japanese immigrants to California, and sons of Japanese immigrants compared with white men in California (aged 45–64 years) to conclude that with immigration, men tended to acquire cancer mortality risks become closer to the host population. While such data exemplify the importance of the role of environment in influencing cancer mortality risks, residual effects of the country of origin (which were significant for the stomach cancer, in particular) do not rule out the possibility of genetic susceptibility difference, or retention of some cultural (specially dietary habits) behaviors of the immigrants even after a generation of migration. More recently, such effects of immigration had been studied in terms of the migration-related adaptation to the culture and health behaviors of the host country population[19,20]. For this type of epidemiological research, it is not enough to group all immigrants into a single class; inter-individual variation of migrants must be measured in a scale of migration history scores, and thus, work described in the present work becomes methodologically relevant. Since our data refers to the formulation and validation of a composite migration history score for Mexican-American women, a brief commentary on recent research on immigration/acculturation-related health behavior changes in Mexican-Americans may exemplify further importance of an appropriate definition of the composite migration history score developed here.

Immigration/acculturation and its effects on health behavior changes in Mexican-Americans

In studying the influence of immigration on changing health behaviors, immigration is generally conceptualized as žacculturation', the process of change as a result of contact between cultural groups[21,22]. One aspect of acculturation is cultural assimilation; i.e., the adoption by minorities of the larger group's cultural norms. In epidemiological studies of Mexican-Americans, acculturation has been measured in terms of one or both of the two dimensions: 1) language preference

(in reading, radio program listening and television viewing) and 2) country of ancestry (country of birth of subjects, their parents and grand parents) and length of residence in the USA[17,23]. Of these the Cuellar Acculturation scale[23] is more extensive and perhaps more rigorous. On the language dimension it constructs a scale based on 11 language-related items, such as first language learned; understanding of spoken and written English; language preference with spouse, children, grand-children and friends; and language preference for reading books and newspapers, watching television and listening to radio programs, and the like. Each item is rated from Spanish/Latino orientation to English/Non-Latino orientation by nominal numeric scores. They are subsequently standardized individually (to a mean of zero and a variance of one).

The other dimension of this acculturation score is based on country of birth of subjects and their immediate ancestors (e.g., parents and grandparents) and years of residency in the USA. Like the language dimension, these itemized responses are also individually standardized, and the grand average of all itemized transformed scores forms the final acculturation score. In applying the Cuellar Acculturation scoring technique in other studies in Mexican-Americans, Markides et al.[17] and Cantero et al.[16] observed that the language acculturation scale exhibits: 1) a strong positive correlation ($r = 0.60$) with subject's generation of residence in the USA; 2) a somewhat weaker but significant positive correlation ($r = 0.42$) with the length of residence in the US for foreign-born immigrants; and 3) a negative correlation ($r = -0.70$) with the age at migration. However, they also noted that the detailed Cuellar Acculturation Scale questionnaire, when administered particularly to the less educated Mexican American women, leaves many of the language items unanswered because of failure of response. Nonetheless, Markides et al.[17] observed that acculturation scored by language preference is highly predictive of health behavior change, in the sense that with increasing trend of adaptation to the language of preference of the larger community, the health habits become closer to the norm of the society with which assimilation occurs[24]. Thus, the validation results emerging from these studies imply that the migration status measured based on birthplaces of subjects and their immediate ancestors, and years of residence in the US may be used as replacement of language-based acculturation scores. In turn such migration-history based scores can also serve the purpose of scaling the immigrant's adaptation ability to the norm of the host population, at least in reference to health behavior. In other words, with a longer history of migration immigrants may be more adapted to the host population norms for a given age set and social and economic class.

As mentioned earlier, the language preference dimension of defining an acculturation scale has several subjective elements. For example, opportunity of exposure to English, education, socio-economic status, etc. affect language preference substantially, and hence, acculturation scale defined by the language preference dimension is not totally independent of these factors, which may themselves have their independent (or moderating/mediating) effects on health behaviors and disease risks. In contrast, as shown in this work, a composite migration history score constructed from the length of residence, birthplaces of subjects and their close relatives is more objective and easier to evaluate through the administration of a simple demographic questionnaire. They are more reliable and objective measurements than language preference, which depends on the

opportunity for exposure to a new language. Birthplace of subjects and their family members (parents, grandparents, etc.) and years of residence in the USA are directly verifiable and readily available demographic variables, with respect to which the exposure to the culture and environment of the host country can be assessed. Furthermore, the adequacy of language preference-based measures of acculturation has been criticized on the ground that general illiteracy may also contribute to language preference patterns, irrespective of the migration history of Mexican-Americans[25]. Particularly for Mexican-American middle-aged and elderly women, Cueller[26] estimated that less than half are not fully literate in either English or Spanish.

In view of these, we conclude that the objective method of developing a composite migration history score based on demographic data on birthplaces of subjects and their immediate ancestors, along with the length of residence in the host country can capture inter-individual variation of migration history. The composite migration history score, thus developed, can be subsequently used along with other variables (such as family stress, socio-economic status, health behaviors, etc.) to investigate whether or not migration *per se* affects health outcomes, or the migration-related health outcomes are mediated by more direct effects of other risk factors.

Acknowledgements

This research was a part of the author's doctoral dissertation project, completed at the University of Texas School of Public Health at Houston, Texas. Dr. W. H. Mueller supervised the work, and Dr. D. M. Hoelscher, and Professor B. E. Quill gave important suggestions that improved the data analyses substantially. Data analyzed in this work was collected during the project »Unidos en Salud: Weight Loss for Mexican Americans« (John P. Foreyt, PhD – PI), funded by NIDDK grant R18-HL47082.

REFERENCES

1. NATIONAL COUNCIL OF LA RAZA: On the sidelines: Hispanic elderly and the continuum of care. (National Council of La Raza, Washington DC, 1991). — 2. VILLARREAL, S. F.: Current issues in Hispanic health. In: Dialogue on health concerns for Hispanics. (US Department of Health and Human Services, Washington DC, 1984). — 3. NATIONAL HEART, LUNG, AND BLOOD INSTITUTE: The fifth report of the Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure: National high blood pressure education program. (National Institutes of Health, Washington DC, 1992). — 4. US DEPARTMENT OF HEALTH AND HUMAN SERVICES: Health status of the disadvanged: Chartbook 1990. (HRSA Publication HRS-P-DV 90-1, Government Printing Office, Washington DC, 1990). — 5. STERN, M. P., Ethnicity Dis., 3 (1993) 7. — 6. BERKMAN, L. F., L. BRESLOW, D. WINGARD, Health practice and mortality risk In: BERKMAN, L. F., L. BRESLOW (Eds.): Health and ways of living: The Alameda County Study. (Oxford University Press, New York, 1983). — 7. DAVIS, M. A., J. M. NEUHAUS, D. J. MORITZ, D. LEIN, J. D. BARCLAY, S. P. MURPHY, Prev. Med., 23 (1994) 369. — 8. STERN, M. P., S. P. GASKILL, C. R. ALLEN, V. GARZA, J. L. GONZALEZ, R. H. WALROP, Am. J. Epidemiol., 113 (1981) 546. — 9. FRIIS, R, G. NANJUNDPPA, T. J. PEDERGAST, M. WELSH, Calif. Publ. Hlth. Rep., 96 (1981) 418. — 10. MARÌN, G., E. J. PEREZ-STABLE, B. V. MARÌN, Am. J. Publ. Hlth., 79 (1989) 196. — 11. BLACK, S. A., K. S. MARKIDES, Am. J. Publ. Hlth., 83 (1993) 890. — 12. GUENDELMAN, S., B. ABRAMS, Am. J. Publ. Hlth., 85 (1995) 20. — 13. GOSLER, P. W., C. A. MACERA, L. G. CASTELLANOS, J. R. HUSSEY, F. S. SY, P. A. SHARPE, Ethnicity Dis., 7 (1997) 106. — 14. MARKIDES, K. S, L. A. RAY, C. A. STROUP-BENHAM, F. TREVIÑO, Am. J. Publ. Hlth., 80 Suppl. (1990) 42. — 15. POSTON, W.S., C. K. HADDOCK, N. E. OLVERA, R. R. SUMINSKI, R. S. REEVES, J. K. DUNN, C. L. HANIS, J. P. FOREYT, Am. J. Hlth. Behav., 25 (2001) 396 — 16. CANTERO, P. J., J. L. RICHARDSON, L. BAEXCONDES-GARBANATI, G. MARKS, Ethnicity. Dis., 9 (1999) 166. — 17. MARKIDES, K. S., N. KRAUSE, C.

F. MENDES DE LEON, Am. J. Publ. Hlth., 78 (1988) 1178. — 18. BUELL, P., J.E. DUNN, Cancer, 18 (1965) 656. — 19. NEFF, J. A., S. K. HOPPE, Alcohol and Alcoholism, 27 (1992) 293. — 20. BALCAZAR, H., F. G. CASTRO, J. L. KRULL, Health Educ. Quart., 22 (1995) 61. — 21. KEEFE, S. E., Acculturation and the extended family among urban Mexican Americans. In: PADILLA, A. M. (Ed.): Acculturation theory, models and some new findings. (Westview Press, Boulder, 1980). — 22. BERRY, J. W., Acculturation as varieties of adaptation. In: PADILLA, A. M. (Ed.): Acculturation theory, models and some new findings. (Westview Press, Boulder, 1980). — 23. CUELLAR, I, L. C. HARRIS, R. JASSO, Hispanic J. Behav. Sci., 2 (1980) 199. — 24. MARÌN, G., Issues in the measurement of acculturation among Hispanics. In: GEISINGER, K. E. (Ed.): Psychological testing of Hispanics. (American Psychological Association, Washington DC, 1992). — 25. LOUE, S.: Gender, ethnicity, and health research. (Kluwer Academic/Plenum Publishers, New York, 1999). — 26. CUELLAR, J. B., Hispanic -American aging: Geriatric education curriculum developed for selected health professionals. In: HARPER, M. S. (Ed.): Minority aging: essential curricula for selected health and allied health professionals. (DHHS Publ. No. HRS P-DV-90-4, US Government Printing Office, Washington DC, 1990).

B. M. Chakraborty

Division of Epidemiology and Biostatistics, Department of Environmental Health, University of Cincinnati, 3223 Eden Avenue, Cincinnati, OH 45267, USA

FORMULACIJA I PROCJENA SLOŽENE VARIJABLE MIGRACIJSKE POVIJESTI ZA EPIDEMIOLOŠKA ISTRAŽIVANJA U HISPANO -AMERIČKIH ŽENA

S A Ž E T A K

U ovom istraživanju razvijena je složena varijabla migracijske povijesti koja uključuje podatke o mjestu rođenja ispitanika, njihovih roditelja i djedova te trajanju boravišta u SAD-u za 390 hispano-američkih žena prekomjerne tjelesne težine. Složena varijabla migracijske povijesti provjerena je faktorskom analizom devet-dimenzionalnih varijabli migracijske povijesti. Važnost ove metodologije definiranja složene varijable migracijske povijesti za epidemiološka istraživanja raspravljena je u kontekstu ispitivane uloge promjena zdravstveno-relevantnih komponenti ponašanja koje su poveze s migracijom te posljedične uloge koju takvo ponašanje ima u riziku zdravstvenih posljedica migranata.