An Indirect Method to Estimate Cancer Incidence Rates for Specific Hispanic/Latino Groups

Holly Howe, PhD
Andrew J. Lake, MS
Brenda K. Edwards, PhD
Maria J. Schymura, PhD
Acknowledgements

- Holly Howe
- Brenda Edwards
- Maria Schymura
U.S. Hispanic/Latino Population Background

• Hispanic/Latinos are the largest and fastest growing minority race/ethnic population in U.S.
Abstract

• Historically, states with large Hispanic populations have been the source for Hispanic cancer rates.

• Aggregation of data limited by different methodologies to identify Hispanic persons.

• Data are now available for more 85% of U.S. Hispanic population.

• NAACCR developed a standard method to enhance the identification of Hispanic persons with cancer – the NAACCR Hispanic/Latino Identification Algorithm, NHIAv2

• State rates suggest regional diversity in their Hispanic cancer profiles.
Hispanic Subgroup Analysis

• Purpose is to examine Hispanic rates among different Hispanic/Latino populations.

• Because of diversity of Hispanic populations, geography would be too crude a measure of cancer rates for specific Hispanic groups.

• Hispanic Latino subgroup population data at the county level are available only in the Census 2000 data.

• How can incidence rates be calculated for Hispanic subgroups?
Analysis Approach

• Use the 1995-2004 CINA Deluxe File

• Group counties with homogenous Hispanic/Latino populations, based on the Census 2000 data.

• Calculate and evaluate Hispanic/Latino incidence rates for each group.
CINA Deluxe File

• Contains Cancer Hispanic/Latino incidence data from 1995-2004.

• Includes data from 34 registries that meet the NAACCR “Fit For Use” criteria.

• Includes county level Incidence data for 1,955 counties.

• Includes the derived Hispanic field NHIAv2.

• Includes data for over 85% of U.S. Hispanic population.
Grouping Counties

- Census 2000 SF2 file contains subgroup populations by county for total U.S.

- For each county, determine the population percentage of each major Hispanic/Latino subgroup.
Census 2000 SF2 Hispanic Categories

Hispanic or Latino (of any race)
  Mexican
  Puerto Rican
  Cuban
  Dominican Republic
  Central American
    Costa Rican
    Guatemalan
  Honduran
  Nicaraguan
  Panamanian
  Salvadoran
Census 2000 SF2 Hispanic Categories

South American
   Argentinian
   Bolivian
   Chilean
   Colombian
   Ecuadorian
   Paraguayan
   Peruvian
   Uruguayan
   Venezuelan

All Other Hispanic/Hispanic NOS
• Census suppresses counts for any subgroup where there are < 100 people.

• For many counties, we were unable to reconcile the total Hispanic/Latino counts with subgroup counts.
• The population counts for Dominicans, South and Central Americans were too small to be a predominant category.

• Narrowed the subgroups into five categories:
  – Mexican
  – Puerto Rican
  – Cuban
  – All Other Hispanic groups (includes Dominicans, South/Central Americans)
  – Hispanic NOS
Initial Observations

- Because of the way Census suppresses data, 2/3 of all counties had suppression for at least one of the 5 groups.

- Suppression did not affect selection of homogenous counties.
Additional Adjustments

- To maximize the number of counties included, the most predominant subgroup was attributed a percentage of the Hispanic NOS population.

- The adjustment was based on the percentage of the largest group and the number of people in the NOS group.
Additional Adjustments

• For example, if 65% of the Hispanic/Latino residents were known to be Mexican, it was assumed that 65% of the Hispanic NOS were also Mexican.

• The 65% NOS counts are added to the majority group – Mexican in this example.
Other County Information

• The CINA Deluxe file contains data for 1,955 counties

• Of these counties, 292 had less than 100 total Hispanic population as defined in the Census 2000 SF2 file.

• Of the remaining eligible counties, 67 counties had no Hispanic cases.

• Data from 1,663 counties were included for this analysis.
Homogeneous Counties

• What is a homogeneous county?

• Homogeneity of county Hispanic population was evaluated using three thresholds:
  > 50%
  > 60%
  > 75%
Homogeneous Counties

• Higher thresholds reduced the potential of misclassification of cases to a specific group.

• Higher thresholds substantially reduce the number of counties and cases included to create meaningful and stable incidence rates.
## Counties Included

<table>
<thead>
<tr>
<th>Hispanic Homogenous Threshold Level</th>
<th>Number of Counties Included</th>
<th>Percent of Total U.S. Hispanic Population 1995-2004</th>
<th>Number of Hispanic/Latino Cases Included CINA Deluxe 1995-2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 50%</td>
<td>1,216</td>
<td>80.8 %</td>
<td>528,592</td>
</tr>
<tr>
<td>&gt; 60%</td>
<td>1,056</td>
<td>68.7 %</td>
<td>401,265</td>
</tr>
<tr>
<td>&gt; 75%</td>
<td>712</td>
<td>64.0 %</td>
<td>369,607</td>
</tr>
</tbody>
</table>
## Counties Included

<table>
<thead>
<tr>
<th>Hispanic Homogenous Threshold Level</th>
<th>Number of Counties Included</th>
<th>Percent of Total U.S. Hispanic Population 1995-2004</th>
<th>Number of Hispanic/Latino Cases Included CINA Deluxe 1995-2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 50%</td>
<td>1,216</td>
<td>80.8 %</td>
<td>528,592</td>
</tr>
<tr>
<td>&gt; 60%</td>
<td>1,056</td>
<td>68.7 %</td>
<td>401,265</td>
</tr>
<tr>
<td>&gt; 75%</td>
<td>712</td>
<td>64.0 %</td>
<td>369,607</td>
</tr>
</tbody>
</table>
## Counties Included

<table>
<thead>
<tr>
<th>Hispanic Subgroup</th>
<th>Number of Counties Included</th>
<th>Number of Hispanic/Latino Cases Included CINA Deluxe 1995-2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexican</td>
<td>1,093</td>
<td>383,582</td>
</tr>
<tr>
<td>Puerto Rican</td>
<td>86</td>
<td>61,780</td>
</tr>
<tr>
<td>Cuban</td>
<td>2</td>
<td>58,085</td>
</tr>
</tbody>
</table>
Hispanic Subgroup Distribution

CINA Deluxe 1995-2004

- Mexican: 383,582 (73%)
- Cuban: 58,085 (11%)
- Puerto Rican: 61,780 (12%)
- Other Hispanic: 23,989 (4%)
- Hispanic NOS: 1,156 (0%)

Total: 525,692
Methods

• Using SEER*Stat, age adjusted Hispanic rates, and 95% gamma confidence intervals were computed.

• Sex–specific Hispanic Rates were computed using NHIAv2 for each of the following:
  – Mexican, Puerto Rican, and Cuban
  – All Hispanics
  – Non-Hispanic Whites.

• Rates were computed for 55 different sites.
Results

• The profiles for the three Hispanic/Latino groups differed from one another.

• In general, U.S. Mexicans have lower cancer incidence rates than the U.S. Hispanic rates, while Puerto Ricans and Cubans are higher than U.S. Hispanic rates.
Incidence Rates For All Sites
Combined by Subgroup and Sex

- Mexican
- Puerto Rican
- Cuban
- Hispanic
- NHW

[Bar chart showing incidence rates by subgroup and sex]
Results: Mexicans

• Rates for Mexican men and women were significantly lower for most cancer sites.

• Exceptions were
  – Kidney for Mexican men
  – Liver, kidney and gallbladder among Mexican Latinas.
Results: Puerto Ricans

• Rates for Puerto Rican men and women were higher than all Hispanics combined, except for:
  – Kidney
  – Testis among males

• Compared to all Hispanics combined, the greatest differences for Puerto Rican Latinas occurred in:
  – Nasopharynx
  – Oropharynx
  – Esophagus
  – Kaposi sarcoma
Results: Puerto Ricans

• Compared to all Hispanics combined, the greatest differences for Puerto Rican men occurred in:
  – Several subsites of the oral cavity, including tongue, nasopharynx, and oropharynx
  – Esophagus
  – Penis
  – Hodgkin and non-Hodgkin lymphomas
  – Kaposi sarcoma
Results: Cubans

- Cuban men and women had many of the highest rates of cancer incidence among the three subgroups.

- Compared to all Hispanics combined, the greatest differences among Cuban men occurred in:
  - Many subsites of oral cavity and pharynx
  - Colon
  - Anus
  - Larynx
  - Lung and bronchus
  - bones and joints
  - Urinary bladder
  - Breast \textit{in situ}
  - Melanoma
  - Kaposi sarcoma
Results: Cubans

- Compared to all Hispanics combined, the greatest differences among Cuban women occurred in:
  - Many subsites of floor and gum of the mouth
  - Colon
  - Anus
  - Breast *in situ*
  - Uterine corpus
  - Bladder
  - Brain
  - Thyroid
  - Hodgkin Lymphoma
  - Leukemia
Results: Cubans

- Cuban men had several sites that were significantly lower than all other Hispanics combined. And the lowest of the three subgroups.
  - Esophagus
  - Stomach
  - Liver
  - Penis
  - Myeloma

- Cuban women had lowest rates for the cancers:
  - Liver
  - Gallbladder
  - Stomach
  - Esophagus
  - Kidney
  - Cervix
Results: Non-Hispanic White

- For the most part, rates in the three subgroups were lower than those of Non-Hispanic white population.

- Cuban rates for men and women were more similar to non-Hispanic White rates than the other subgroups.
Results: Non-Hispanic White

• For some sites, rates for Cubans (male and female) were higher than rates for non-Hispanic whites.
  – Several subsites of oral cavity
  – Stomach
  – Colon
  – Anus
  – Liver
  – Gallbladder
  – Larynx (men)
  – Cervix
  – Bones and joints
  – Kaposi sarcoma
Results: Non-Hispanic White

• Compared to Non-Hispanic white population, Mexican men had higher rates for:
  – Stomach
  – Liver
  – Gallbladder
  – Penis
  – Kaposi sarcoma
Results: Non-Hispanic White

• Compared to Non-Hispanic white population, Mexican Latinas had higher rates for:
  – Stomach
  – Liver
  – Gallbladder
  – Cervix
  – Kidney
  – Myeloma
  – Kaposi sarcoma
Results: Non-Hispanic White

- Compared to Non-Hispanic white population, Puerto Rican men had higher rates for:
  - Oral Cavity and Pharynx
  - Stomach
  - Esophagus
  - Liver
  - Gallbladder
  - Larynx
  - Prostate
  - Penis
  - Hodgkin and non-Hodgkin Lymphoma
Results: Non-Hispanic White

• Compared to Non-Hispanic white population, Puerto Rican Latina had higher rates for:
  – Nasopharynx
  – Esophagus
  – Stomach
  – Liver
  – Gallbladder
  – Cervix
  – Uterine NOS
  – Myeloma
  – Kaposi sarcoma
Limitations

• It would be preferable to have Hispanic heritage recorded in the medical record.

• Annual counts of specific Hispanic ethnic populations by sex and age are needed for rates calculations.

• Misclassification of specific Hispanic groups could dilute any real differences among Hispanic subgroups.

• Some differences in cancer rates could be attributed to regional differences in cancer rates themselves.
Conclusions

• Indirect approach to creating age adjusted cancer rates for specific Hispanic/Latino subgroups provides useful information for exploring the differences among populations as well as developing relevant cancer control interventions.
Conclusions

• While benefits of this method outweigh the limitations, it would be preferable to have direct identification of specific ethnicity in the medical record, and annual population estimates given that the Hispanic population is the second largest race-ethnic group in the U.S.