1. Introduction

The purpose of the Oregon State Health Profile Report is to provide a picture of the health of the Oregon population as a whole. However, it is also important to know whether there are variations in health status in different populations or regions of the state. Below, we present analyses of 31 of the State Health Profile indicators by:

- Ethnicity and race, and
- Coordinated Care Organization (CCO) region

We chose ethnicity and race because of the significant health disparities that exist across these populations in Oregon and in the nation as a whole. It is important to highlight and track these disparities to assess whether progress is being made to close these gaps.

We also assessed health status within each of the 16 CCO regions to provide information on baseline health status as the CCOs begin a process of assessment and planning for their regions.

It is important to note that the CCO analyses shown here represent health status among the residents of each CCO region, not CCO members. Also, many CCOs boundaries overlap, and persons who reside in overlapping areas were counted as residents of any CCO serving that area. This limits comparisons of one CCO to another because many comparisons will not be based on mutually exclusive groups. However, we have provided a state estimate for each indicator so that estimates for CCO could be compared to the state overall.
The indicators are broad: they include measures of mortality and morbidity, premature death, health risk behaviors among adults and teens, and measures of maternal and child health. They represent a subset of the indicators used for the State Profile because they are confined to those data sources that lie within public health, and could be analyzed by ethnicity-race and CCO. However, taken together, they are sufficient to provide an overall picture of the health status of each of the populations analyzed.

2. Analyses by Coordinated Care Organization (CCO): Overview

There are currently 16 CCOs in Oregon. In many cases, CCO regions comprise one or more counties and county-level analyses could be used to describe CCO population indicators. However, in some cases boundaries of a CCO extended over county lines. The use of zip codes to define the limits of each CCO service area allowed us to estimate population for each CCO, even for those whose boundaries did not correspond to county boundaries.

As noted, these analyses focus on the population residing within the CCO region, not on the patient population of a particular CCO. Because several of the CCOs overlap in their service regions, we counted persons residing in overlapping CCO regions as part of the service area for any CCO that provided service in that region. Therefore, the sum of the population of persons residing within all the CCO regions combined will be greater than the state population.

The overlapping nature of the CCO regions also makes it difficult to conduct direct comparisons across CCOs, as many CCO regions share population. However, state-level estimates (both unadjusted and age-adjusted) for each indicator are provided for CCO-total state comparisons.

2.1. Calculation of CCO denominators

2.1.1. Survey Data (Behavioral Risk Factor Surveillance System (BRFSS), Oregon Health Teens (OHT),

Calculating the survey sample within each CCO was done using zip code of the BRFSS survey respondent. For analyses of OHT data, zip code of the school was used to determine CCO region.

2.1.2. Population-based data (Vital Statistics, Cancer Registry, Hospital Discharge Index, Orpheus, Alert)

We developed an estimation procedure to calculate population within a Coordinated Care Organization (CCO) geographic region. Although CCOs are defined using zip code boundaries, there are no population files that provide data by zip code, age and sex. So, we relied on the census projected population files for the year needed that provide information by county, age and sex. To the extent that a CCO was comprised of one or more counties in their entirety, population estimates were straightforward. However, several CCO boundaries crossed county lines. In these cases, we identified the zip codes that lay within each of the counties comprising the CCO. For each group of zip codes that belonged to a particular county, we then used the 2010 Zip Code Tabulation Areas (ZCTA) file produced by the U.S. Census bureau for 2010 to sum the populations across those zip codes. We then divided this sum by the total population within each county. This yielded an estimate of the proportion of the county’s population that belonged to a particular CCO region. These fractions were then applied to the total census-
derived projected county population data for the appropriate year (and subgroup, if needed, eg. age or sex), in order to estimate the total population for the CCO.

**Example:** Primary Health of Josephine County is a CCO that contains all of Josephine County, part of Jackson County, and part of Douglas County. We separately summed the population of each zip code for this CCO that lies within Jackson and Douglas counties using the ZCTA file and then calculated those quantities as proportions of the Jackson or Douglas county populations equally by age and/or sex, if needed for each county. Those proportions were applied to the total population count for each county for the appropriate year or group of years to arrive at the population denominator for the CCO.

One limitation in the method lies in the assumption that the contribution of a group of zip codes to a county population is roughly equivalent when considering the total population or considering a sub-group, such as women of childbearing age. This assumption may be violated for certain sub-populations (such as persons of a particular age), as they may be concentrated in particular geographic regions.

### 2.1.2.1. Population data sources

We used two sources for population data. U.S. Census Bureau Population estimates\(^1\) were used when indicator data included a multiracial category. If the database for a particular indicator did not include a multiracial category, we used population files from the National Center for Health Statistics.\(^2\) Additional information on uses of these files is included below on the section on calculation of denominators by ethnicity and race.

### 2.1.2.2. Zip codes that cross county lines

There were a few instances where a particular zip code crossed the county line. In these situations we used a variation of the same methods. We estimated the proportion of the total county population represented by the area of the zip code that extends into the adjacent county (data also available from ZCTA files) and applied that proportion according to the methods outlined above.

### 2.1.2.3. Estimation of numerator (event) data by CCO when data are only available by county

For numerator data for which county of residence but not zip code was available in the data set, we applied the same proportions to the events occurring in each county. This was only necessary for the HIV new diagnosis rate. In the example given above, we would add 100% of the events from Josephine to the appropriate percent of the events occurring in Jackson, determined by the relative contribution of the Jackson county zip codes that belong to the CCO service area.

This method assumes that the distribution of events is uniform across all zip codes of a county, an assumption that will likely be violated for counties with large population centers. However, only one database required us to use this method (HIV diagnosis rates).

---

\(^1\) Census Bureau Population Estimates: U.S. Census Bureau, Population Division. Annual estimates of the resident population by age, sex, race and Hispanic origin for counties, Vintage 2012.

2.1.2.4. Missing data for residence in a CCO region

If a case was missing on county, on zip code, or had an invalid zip code-county combination we considered that case missing with respect to CCO region. Numbers of missing cases were listed for each table.

3. Analyses by ethnicity and race: Overview

In all databases, ethnicity (Hispanic or non-Hispanic) and race (White, African-American, Asian, Pacific Islander, and American Indian/Alaska Native) are collected separately. For analysis, we constructed a single, mutually exclusive ethnicity-race category by including Hispanic ethnicity as a category, and designating all other categories as non-Hispanic. In this scheme, Hispanic ethnicity was used as an overriding category within the race classification system. For example, if a respondent indicated that they were Hispanic and also indicated they were White, they would be counted under the Hispanic category and removed from the White category.

Some data sources collect information separately on Asians and Pacific Islanders, while others report this information as a combined category. We reported information separately when possible, even if data were sparse for one of the categories.

Some data sources allow for a multiracial classification.

The resulting classifications in the integrated ethnicity-race scheme were: Non-Hispanic White, non-Hispanic African-American, non-Hispanic Asian, non-Hispanic Pacific Islander, non-Hispanic American Indian/Alaska Native, non-Hispanic multiracial, and Hispanic.

3.1. Calculation of denominators by ethnicity and race

The creation of population denominators for the ethnicity-race analyses depended on the classification scheme used by each database. If the database did not allow for a multiracial classification, then NCHS population files were used. All ethnicity-race analyses were conducted for the state as a whole, and thus there were no special procedures needed, as there were for CCO.

Analyses of these and other indicators are available through the Oregon Public Health Assessment Tool (OPHAT). Note however, that the age adjustment categories are different, and thus estimates may not correspond precisely.

3.2. Missing data on ethnicity or race

As noted above, if a case was classified as Hispanic on ethnicity, they were designated as Hispanic, regardless of their classification on race, even if missing on race. However, if a case was missing on ethnicity, but was classified on race, the case was considered missing on the overall ethnicity-race variable as the case could not be classified as “non-Hispanic” within the racial categories. Numbers of cases with missing data on ethnicity-race were noted for each indicator.

4. Statistical considerations: Overview

For all estimates by CCO region, we presented the margin of error (half-width of the 95% confidence interval (CI)) calculated by using the normal approximation to the binomial distribution (1.96*standard error).
For estimates by ethnicity-race, we presented the standard error only. In both cases, statistics were calculated only for the age-adjusted estimates (see below).

Standard errors for age-adjusted estimates were calculated in Excel, using the following formula:

\[ SE_{AAR} = AAR \times \sqrt{\frac{\sum w_i^2 \times r_i^2}{\sum [w_i \times r_i]}} \]

Where \( w \)=weight, \( r \)=rate, \( n \)=event count, and \( i \)=age-groupings used for adjustment.

4.1. Identification of Disparities

We used statistical testing to determine if the estimate for each racial or ethnic group was different from the estimate for non-Hispanic Whites. We tested the equality of means at a significance level of 0.05. We chose to conduct these tests on the age-adjusted estimates.

It should be noted that full determination of health disparities in a population should generally not rely on statistical differences alone. Magnitude of the difference, trends over time, and national patterns should also be considered.

4.2. Age adjustment

We used the U.S. 2000 population as a reference population. In general, we used three age categories for the standardization procedure: 18-34, 35-54, 55 and over. Some indicators were defined within a specific age range and age standardization was unnecessary (as in teen pregnancy rates among girls aged 15-17). Several other indicators were age restricted, but still required adjustment. For falls among adults 65 and older, we used two categories for adjustment: 65-74 and 75 and over. For percent of births that are low birthweight and percent of women initiating prenatal care in the first trimester, we used three categories: 24 and under, 25-34 and 35 and over.

The statewide value for each indicator was adjusted to the distribution of the U.S. population for the year 2000.

4.3. Combining multiple years of data

For some indicators counts were sparse. As a general rule, if there were fewer than 2000 events for a particular indicator we combined years of data in 3 year or 5 year increments until we reached 2000. We did not use more than 5 years of data for any indicator.

4.4. Guidelines used for small numbers

4.4.1. Confidentiality rules

In census-based analyses such as those from birth or death certificates or population registries, we followed a general guideline that suppresses estimates based on a denominator of fewer than 50 persons in order to protect against a potential breach of confidentiality. However, none of the estimates in this report needed to be suppressed, as all had denominators over 50.
For survey data, we assumed that denominator data are not identifiable, as they come from a random sample of a large population and thus confidentiality rules do not apply. One exception to this general guideline concerns data from Oregon Healthy Teens. In the case of geographic analyses, such as those by CCO, a sampled school could be identifiable because it was the only school in the CCO region, or the only participating school among several sampled. In this case, the guidelines dictate that we treat the data as population data rather than survey data, and suppress any estimate based on fewer than 50 students. There were no such schools in our sample.

4.4.2. Reliability rules

We used a general guideline based on relative standard error (RSE) to flag estimates that should be considered statistically unreliable. The RSE is calculated by dividing the standard error of the estimate by the rate or proportion in question. For both population data (rates) and survey data we flagged any estimate with an RSE of greater than 30%. We did not suppress any estimate based on RSE alone.

**Population-based data (rates)**

For crude incidence and mortality rates, the RSE is calculated as follows:

\[
\text{Relative standard error} = \frac{\text{standard error}}{\text{rate}} \times 100
\]

For crude rates, the standard error is calculated as follows:

\[
SE = \frac{\text{rate}}{\sqrt{\text{cases}}}
\]

so the relative standard error is:

\[
\text{RSE} = \frac{\text{rate}}{\sqrt{\text{cases}}} \times \frac{1}{\text{rate}} \times 100 = \frac{1}{\sqrt{\text{cases}}} \times 100
\]

Using a simple rule-of-thumb based on the above formula we suppressed unadjusted rates based on 11 cases or fewer. For age-adjusted rates, we calculated the ratio of the standard error to the rate.

**Survey data (proportions)**

For survey data we generated the RSE by dividing the standard error as calculated by SAS (accounting for the complex survey design in BRFSS and OHT), by the proportion for the indicator in question.

**Additional methods notes by indicator are available from:**

Barbara Pizacani, PhD  
Program Design and Evaluation Services  
Oregon Public Health Division and Multnomah County Health Department  
barbara.a.pizacani@state.or.us