Research article

Calculating expected years of life lost for assessing local ethnic disparities in causes of premature death

Tomás J Aragón*1,2,3, Daphne Y Lichtensztajn2, Brian S Katcher2, Randy Reiter2 and Mitchell H Katz2,3

Address: 1Division of Epidemiology, School of Public Health, University of California, Berkeley, California, USA, 2Community Health Epidemiology, San Francisco Department of Public Health, San Francisco, California, USA and 3Department of Epidemiology and Biostatistics, University of California, San Francisco, California, USA

Email: Tomás J Aragón* - aragon@berkeley.edu; Daphne Y Lichtensztajn - daphne@stanfordalumni.org; Brian S Katcher - brian.katcher@sfdph.org; Randy Reiter - randy.reiter@sfdph.org; Mitchell H Katz - mitch.katz@sfdph.org

* Corresponding author

Abstract

Background: A core function of local health departments is to conduct health assessments. The analysis of death certificates provides information on diseases, conditions, and injuries that are likely to cause death – an important outcome indicator of population health. The expected years of life lost (YLL) measure is a valid, stand-alone measure for identifying and ranking the underlying causes of premature death. The purpose of this study was to rank the leading causes of premature death among San Francisco residents, and to share detailed methods so that these analyses can be used in other local health jurisdictions.

Methods: Using death registry data and population estimates for San Francisco deaths in 2003–2004, we calculated the number of deaths, YLL, and age-standardized YLL rates (ASYRs). The results were stratified by sex, ethnicity, and underlying cause of death. The YLL values were used to rank the leading causes of premature death for men and women, and by ethnicity.

Results: In the years 2003–2004, 6312 men died (73,627 years of life lost), and 5726 women died (51,194 years of life lost). The ASYR for men was 65% higher compared to the ASYR for women (8971.1 vs. 5438.6 per 100,000 persons per year). The leading causes of premature deaths are those with the largest average YLLs and are largely preventable. Among men, these were HIV/AIDS, suicide, drug overdose, homicide, and alcohol use disorder; and among women, these were lung cancer, breast cancer, hypertensive heart disease, colon cancer, and diabetes mellitus. A large health disparity exists between African Americans and other ethnic groups: African American age-adjusted overall and cause-specific YLL rates were higher, especially for homicide among men. Except for homicide among Latino men, Latinos and Asians have comparable or lower YLL rates among the leading causes of death compared to whites.

Conclusion: Local death registry data can be used to measure, rank, and monitor the leading causes of premature death, and to measure and monitor ethnic health disparities.
Background

A core function of local health departments is to conduct public health surveillance, including population health assessments [1,2]. Public health surveillance is the ongoing, systematic collection, analysis, interpretation, and dissemination of data regarding a health-related event for use in public health action to reduce morbidity and mortality and to improve health [3]. For a local health jurisdiction, primary data collection, such as representative population-based surveys, can be expensive and unsustainable. Therefore, local health departments must analyze existing health data, preferably those that are population-based, comprehensive, readily available, and locally relevant.

Death records are an important data source for assessing population health and health disparities because they cover the whole population and include information on key characteristics of decedents, including age, sex, ethnicity, place of residence and of death, and underlying and contributing causes of death. However, cause-specific mortality is typically reported using traditional epidemiologic measures, especially counts and rates (including age-adjusted rates), that are heavily influenced by deaths among older residents. For most causes, these measures are not very sensitive to deaths occurring at younger ages, which are more likely to be premature, preventable deaths.

To identify and prioritize causes of premature death, the standard expected years of life lost (YLL) measure, as developed by the Global Burden of Disease Study [4], provides a valuable analytic tool which can be applied to local geographic areas. The YLL is based on comparing the age of death to an external standard life expectancy curve, and can incorporate time discounting and age weighting. The YLL, combined with the years lived with disability (YLD) measure, make up the disability-adjusted life year (DALY). Unfortunately, directly measuring YLDs (and therefore DALYs) is cost prohibitive and not practical for most local health jurisdictions. In contrast, YLLs are measurable for a comprehensive set of conditions. YLLs, as opposed to more traditional mortality measures (counts, rates, etc.), highlight premature deaths. Where population estimates are available, age-adjusted YLL rates allow comparisons across groups or over time [5]. These deaths are particularly important from a public health and public policy perspective because they represent preventable loss of life.

Although the YLL is a valid, stand-alone measure for identifying and ranking the causes of premature death for a region [6-8], this measure has not been widely adopted for local area mortality analyses. There are several reasons for this. First, detailed methods for calculating YLL are not available in standard epidemiology textbooks or scientific journal articles. In contrast, the years of potential life lost (YPLL) is commonly used: it is easily calculated by subtracting the age of death from a chosen cut-off (e.g., 65, 75, or 85 years) [9,10]; however, the YPLL does not measure deaths after the cut-off age, and it does not incorporate time-based discounting used in cost-effectiveness analysis. The YLL measures every death and can incorporate discounting. Second, with few exceptions [5], sufficient local area YLL analyses have not been published to demonstrate their value in assessing population health. And third, readily available software solutions to make analyses more efficient have not been developed.

The purpose of this paper is to provide detailed methods for calculating YLL for a local geographic area (San Francisco, California, United States), and to demonstrate its value as a population health measure to impact local public health priorities. We illustrate how to use YLL to rank causes of death, and how to use average YLLs to identify the leading causes of premature death for major ethnic groups. Analysis and interpretation of death registry data using YLLs provide objective evidence for public health policymakers, partners, and stakeholders to inform and guide setting of local public health priorities. This is especially important because of the geographic and demographic variation in health outcomes, major risk factors, and health disparities [5,11].

Methods

Overview

Summarized in Table 1 are the notation and definitions used in this article. For the years 2003–2004, registered deaths for San Francisco were obtained from the State of California, Department of Health Services, Center for Health Statistics [12]. The data file contained the underlying causes of death of San Francisco residents (whether or not they died in San Francisco), and the underlying causes of death of non-residents that died in San Francisco. For this study, only San Francisco resident underlying causes of death were used. Population estimates were obtained from the State of California, Department of Finance, Demographic Research Unit [13]. Standard life expectancy figures for men and women are from the Coale-Demeny Model Life Tables West: Levels 25 for men, and Level 26 for women (Table 2). For calculating age-standardized rates we used the Year 2000 United States standard million population (Table 3).

The deaths and population estimates were aggregated into 19 age intervals (columns 1 and 2 of Table 4). For age-standardizations, 11 age intervals were used (Tables 3 and 5). The 19-level age intervals were used for calculating expected years of life lost (YLL) for men and women, (Table 4), stratified by sex and ethnicity (Table 6), strati-
Table 1: Summary of notation and definitions

| Notation     | Definition                                                                 |
|--------------|-----------------------------------------------------------------------------|
| Introduced in Equation 1 |                                                                                           |
| $x$          | Age                                                                                       |
| $n$          | Age interval length                                                                   |
| $e_x$        | Life expectancy at age $x$                                                              |
| $e^*_x$      | Standard life expectancy at age $x$                                                     |
| $n^*_x$      | Standard life expectancy for age interval $x$ to $x + n$                               |
| $\bar{a}_x$  | Average age of death for age interval $x$ to $x + n$                                   |
| Introduced in Equations 2 to 5 |                                                                                           |
| $nY_x$       | Expected years of life lost for age interval $x$ to $x + n$                            |
| $nD_x$       | Number of deaths in age interval $x$ to $x + n$                                        |
| $r$          | Discount rate (usually set to 0.03)                                                    |
| $K$          | Age-weighting modulation constant ($K = 0$, no weighting; $K = 1$, weighting)           |
| $\beta$      | Age-weighting constant (set to $\beta = 0.04$)                                        |
| $C$          | Adjustment constant for age-weights (set to 0.1658)                                    |
| $YLL$        | Expected years of life lost ($YLL = \sum_{x} nY_x$)                                    |
| Equation 6   |                                                                                           |
| ASYR         | Age-standardized YLL rate ($\sum_{x} nY_x^*$)                                           |
| $nY_x^*$     | Weighted expected years of life rate for age interval $x$ to $x + n$                    |
| $nN_x$       | Expected years of life lost rate for age interval $x$ to $x + n$                       |
| $nN_x$       | Population (person-year) estimate for age interval $x$ to $x + n$                      |
| $nW_x$       | Year 2000 United States standard population as weights ($\sum_{x} W_x = 1$)             |
| $nW_x$       | Year 2000 United States standard population ($\sum_{x} W_x = 1,000,000$)                |

Table 2: Standard life expectancies based on Model Life Table West, Level 25 and 26

| Age ($x$) | Level 25: Male life expectancy ($e^*_x$) | Level 26: Female life expectancy ($e^*_x$) |
|-----------|----------------------------------------|----------------------------------------|
| 0         | 80.000                                 | 82.500                                  |
| 1         | 79.358                                 | 81.840                                  |
| 5         | 75.383                                 | 77.950                                  |
| 10        | 70.400                                 | 72.990                                  |
| 15        | 65.414                                 | 68.020                                  |
| 20        | 60.438                                 | 63.080                                  |
| 25        | 55.471                                 | 58.170                                  |
| 30        | 50.512                                 | 53.270                                  |
| 35        | 45.565                                 | 48.380                                  |
| 40        | 40.641                                 | 43.530                                  |
| 45        | 35.766                                 | 38.720                                  |
| 50        | 30.990                                 | 33.990                                  |
| 55        | 26.322                                 | 29.370                                  |
| 60        | 21.810                                 | 24.830                                  |
| 65        | 17.499                                 | 20.440                                  |
| 70        | 13.577                                 | 16.200                                  |
| 75        | 10.166                                 | 12.280                                  |
| 80        | 7.447                                  | 8.900                                   |
| 85        | 5.238                                  | 6.220                                   |
| 90        | 3.544                                  | 4.250                                   |
| 95        | 2.311                                  | 2.890                                   |
fied by cause of death and sex (Table 7), and stratified by cause of death, sex, and ethnicity (Additional file 1: Tables A-1 to A-4).

**Cause of death categories**

Using the International Classification of Diseases, 10th Revision (ICD-10) [14], the cause of death categories were adapted from the World Health Organization Global Burden of Disease Study [15] and the Centers for Disease Control and Prevention External Cause of Injury Mortality Matrix [16]. The cause of death category definitions were comprehensive, mutually exclusive, and sufficiently specific to support public health interventions [see Additional file 2]. For example, we used specific cancer diagnoses (e.g., “lung cancer”) instead of broad categories (e.g., “all cancers”).

**Interpolating standard model life table**

For a single death at age x, the YLL for that individual is simply the expected years of life remaining at the age of death (i.e., life expectancy at age x: \( e_x \)) based on the model life table [West, Level 25 and 26 (Table 2)]. However, the table does not contain life expectancies for deaths within age intervals. For ages that fall within an interval, the life expectancy must be interpolated from the table.

For a group of deaths that occurred at ages within age interval \( x \) to \( x + n \), (i.e., \( n = \) age interval length), the expected years of life remaining for those deaths \( (\Delta e_x) \) is estimated using a formula for linear interpolation (Equation 1) [17]:

\[
\Delta e_x = e_x + \left( n \cdot a_x - x \right) \left( e_x + n - e_x \right) \\
\left( x + n - x \right)
\]

where \( a_x \) is the average age of death, and \( e_x \) and \( e_{x+n} \) are Table 2 model life expectancies at ages x and x + n, respectively.

### Table 3: Year 2000 United States standard million population

| Age interval | \( n \) | \( W_x \) | \( nW_x \) | \( \Delta D_x \) | \( \Delta G_x \) | \( g_x \) | \( Y_x \) |
|--------------|--------|----------|----------|----------------|----------------|-------|--------|
| < 1          | 1      | 13,818   | 0.013818 | 28             | 0.1            | 80.0  | 848.6  |
| 1–4          | 4      | 55,317   | 0.055317 | 10              | 2.4            | 78.0  | 301.2  |
| 5–9          | 5      | 145,565  | 0.145565 | 3               | 6.4            | 74.0  | 89.1   |
| 10–14        | 5      | 138,646  | 0.138646 | 4               | 13.3           | 67.1  | 115.5  |
| 15–19        | 5      | 135,573  | 0.135573 | 27              | 17.9           | 62.5  | 762.1  |
| 20–24        | 5      | 126,246  | 0.126246 | 56              | 22.6           | 57.9  | 1,538.0|
| 25–29        | 5      | 127,423  | 0.127423 | 72              | 27.4           | 53.1  | 1,912.6|
| 30–34        | 5      | 111,300  | 0.111300 | 111             | 33.0           | 47.5  | 2,810.4|
| 35–39        | 5      | 156,377  | 0.156377 | 156             | 37.7           | 42.9  | 3,764.2|
| 40–44        | 5      | 249,426  | 0.249426 | 249             | 42.6           | 38.1  | 5,655.8|
| 45–49        | 5      | 325,476  | 0.325476 | 325             | 47.6           | 33.3  | 6,844.8|
| 50–54        | 5      | 434,525  | 0.434525 | 434             | 52.5           | 28.6  | 8,340.5|
| 55–59        | 5      | 454,575  | 0.454575 | 454             | 57.5           | 24.1  | 7,782.5|
| 60–64        | 5      | 403,624  | 0.403624 | 403             | 62.4           | 19.7  | 5,999.3|
| 65–69        | 5      | 465,677  | 0.465677 | 465             | 67.7           | 15.4  | 5,738.5|
| 70–74        | 5      | 547,726  | 0.547726 | 547             | 72.6           | 11.8  | 5,432.1|
| 75–79        | 5      | 787,776  | 0.787776 | 787             | 77.6           | 8.7   | 6,054.1|
| 80–84        | 5      | 887,825  | 0.887825 | 887             | 82.5           | 6.5   | 5,117.3|
| 85+          | 10     | 1,294,90  | 1.29490  | 1,294           | 90.3           | 3.7   | 4,520.3|

| Total        | 95     | 6,312 (Deaths) | 73,626.8 (YLL) | 5,726 (Deaths) | 51,194.2 (YLL) |

**Table 4: Calculating expected years of life lost for San Francisco men and women, 2003–2004**

| Age interval | \( n \) | \( nD_x \) | \( nG_x \) | \( nW_x \) | \( nY_x \) |
|--------------|--------|----------|----------|----------|----------|
| < 1          | 1      | 13,818   | 0.013818 | 848.6    | 915.6    |
| 1–4          | 4      | 55,317   | 0.055317 | 301.2    | 90.8     |
| 5–9          | 5      | 145,565  | 0.145565 | 89.1     | 148.4    |
| 10–14        | 5      | 138,646  | 0.138646 | 115.5    | 58.5     |
| 15–19        | 5      | 135,573  | 0.135573 | 762.1    | 171.8    |
| 20–24        | 5      | 126,246  | 0.126246 | 1,538.0  | 446.5    |
| 25–29        | 5      | 127,423  | 0.127423 | 1,912.6  | 732.8    |
| 30–34        | 5      | 111,300  | 0.111300 | 2,810.4  | 833.6    |
| 35–39        | 5      | 156,377  | 0.156377 | 3,764.2  | 1,224.9  |
| 40–44        | 5      | 249,426  | 0.249426 | 5,655.8  | 1,859.5  |
| 45–49        | 5      | 325,476  | 0.325476 | 6,844.8  | 3,113.3  |
| 50–54        | 5      | 434,525  | 0.434525 | 8,340.5  | 3,833.6  |
| 55–59        | 5      | 454,575  | 0.454575 | 7,782.5  | 4,022.4  |
| 60–64        | 5      | 403,624  | 0.403624 | 5,999.3  | 3,534.3  |
| 65–69        | 5      | 465,677  | 0.465677 | 5,738.5  | 2,683.8  |
| 70–74        | 5      | 547,726  | 0.547726 | 5,432.1  | 4,653.9  |
| 75–79        | 5      | 787,776  | 0.787776 | 6,054.1  | 2,683.8  |
| 80–84        | 5      | 887,825  | 0.887825 | 5,117.3  | 6,385.5  |
| 85+          | 10     | 1,294,90  | 1.29490  | 4,520.3  | 9,200.3  |

Total 95 6,312 (Deaths) 73,626.8 (YLL) 5,726 (Deaths) 51,194.2 (YLL)
tively. See Table 4 for use of $n_x e_x$ in spreadsheet calculations.

Calculating expected years of life lost ($Y_{LL}$)

For a group of deaths that occurred at ages within age interval $x$ to $x + n$, the crude expected years of life lost is

$$n_x Y_x = (n_x D_x)(n_x e_x)$$

where $n_x D_x$ is the number of deaths between age $x$ and age $x + n$.

To incorporate discounting and age weighting, one would use Equation 3:

$$n_x Y_x = (n_x D_x)\frac{\sum \left(n_x w_x \left(e^{-(r + \beta)x} - e^{-(r + \beta)nx}\right) - \left[-(r + \beta)w_x\right] - \left[-(r + \beta)nx\right] - \left[-(r + \beta)nx\right]}{1 - e^{-(r + \beta)x}}$$

where $z = -(r + \beta)(n_x e_x + n_x a_x)$. For this equation, $r$ is the discount rate, and $\beta$, $C$, and $K$ are age weighting constants (see Table 1 for complete definitions). To include age weighting, $K$ (the modulation constant) can be set to 1. For this study, age weighting was not used ($K = 0$) and $r = 0.03$.

When the discount rate ($r$) is 0, Equation 3 simplifies to Equation 4:

$$n_x Y_x = (n_x D_x)\frac{\sum \left(n_x w_x \left(e^{-(r + \beta)x} - e^{-(r + \beta)nx}\right) - \left[-(r + \beta)w_x\right] - \left[-(r + \beta)nx\right] - \left[-(r + \beta)nx\right]}{1 - e^{-(r + \beta)x}}$$

First, we calculated the expected of years of life lost, comparing men to women, by summing $n_x Y_x$ for all age intervals (Table 4):

$$Y_{LL} = \sum n_x Y_x$$

Using this approach, we calculated $Y_{LL}$s for 117 specific causes of death stratified by sex, and stratified by sex and ethnicity.

Calculating age-standardized expected years of life lost rates

Using the direct method [18], we calculated age standardized $Y_{LL}$ rates (ASYR). First, we calculated age-specific rates of years of life lost ($n_x Y_x$). Then, these rates were reweighted using using the Year 2000 United States standard million population ($n_x w_x$ in Table 3) [18]. The reweighted rates ($n_x Y_x$) were summed to get an ASYR (Equation 6).

$$ASYR = \sum \left(\frac{n_x Y_x}{n_x N_x}\right)$$

See Table 5 for use of Equation 6 in spreadsheet calculations.

Ranking leading causes of premature death

Determining the leading causes of premature death required two steps. First, the leading 15 causes of death
were ranked by YLL values for San Francisco (stratified by sex), and the leading 10 causes of death were ranked for each ethnic group (stratified by sex). Second, to highlight conditions of highly premature death, these were further subranked by their average YLL values. Age-standardized YLL rates were included to allow comparisons of ethnic groups within sex strata.

**Numerical computing**

All analyses and graphics were conducted in R – a widely available, open source programming language for statistical computing and graphics [19]. To facilitate the YLL calculation for readers, we provide and demonstrate a numerical function for R [see Additional file 3].

**Results**

Displayed in Table 4 is the spreadsheet format for calculating expected years of life lost (YLL) for San Francisco men and women. The age-interval specific number of deaths ($nD_x$), average age of death ($a_x$), standard life expectancy ($e_x^s$), and years of life lost ($Y_x$) are shown. In the years 2003–2004, 6312 men died with 73,627 years of life lost, and 5726 women died with 51,194 years of life lost.

Displayed in Table 5 is the spreadsheet format for calculating direct age-standardized YLL rates (ASYR) for San Francisco men and women, combining years 2003–2004. The sex and age-interval specific population estimates ($nN_{x}$), expected years of life lost ($Y_x$), expected YLL rate ($nY_{x}$), and weighted expected YLL rate ($nY_{x}^{w}$) are displayed in each column. The ASYR for men was 65% higher compared to the ASYR for women (8971.1 per 100,000 persons per year vs. 5438.6 per 100,000 persons per year). Displayed in Table 6 are the YLL, number of deaths, average YLL, age-standardized YLL rates, and ASYR ratios stratified by sex and ethnicity. While whites and Asians account for the largest number of deaths (as expected based on population estimates), African American men and women have the highest age-standardized YLL rates. For all causes of death, the ASYR for African American men is 2.44 times higher compared to white men, and the ASYR for African American women is 2.31 times higher compared to white women.

The leading causes of premature death for San Francisco residents, ranked first by YLLs and then subranked by
average YLLs, are displayed in Table 7. Among the top fifteen, the leading causes with the greatest degree of prematurity of premature deaths are those with the largest average YLLs. Therefore, among men, the leading causes of premature deaths were HIV/AIDS (average YLL = 20.3 years), suicide (19.9 years), drug overdose (21.7 years), homicide (25.0 years), and alcohol use disorder (17.4 years). The leading causes of deaths with the smallest average YLLs were ischemic heart disease (8.9 years), lung cancer (10.7 years), stroke (8.2 years), hypertensive heart disease (11.8 years), and chronic obstructive pulmonary disease (8.3 years).

Among women, the leading causes of premature deaths (those with the largest YLLs) were lung cancer (10.4 years), breast cancer (13.4 years), hypertensive heart disease (8.2 years), colon cancer (9.2 years), and diabetes mellitus (8.6 years). The leading causes of deaths with the smallest YLLs were ischemic heart disease (6.6 years), stroke (6.9 years), chronic obstructive pulmonary disease (7.8 years), pneumonia (5.6 years), and dementias (4.6 years).

Similarly, an analysis was conducted to rank the leading causes of premature death by ethnicity and sex: African Americans (Table A-1), Asians/Pacific Islanders (Table A-2), Latino/Hispanics (Table A-3), and whites (Table A-3). Similar analyses were done for each ethnic group. For example, among African American men, the leading causes of premature death (largest average YLLs) were homicide (25.9 years), HIV/AIDS (19.7 years), hypertensive heart disease (14.7 years), drug overdose (21.7 years), and suicide (20.3 years).
dose (19.8 years), and alcohol use disorder (15.6 years). The leading causes of death with the smallest average YLLs were ischemic heart disease (11.8 years), lung cancer (13.1 years), stroke (10.7 years), chronic obstructive pulmonary disease (11.8 years), and diabetes mellitus (12.6 years).

Age-standardized YLL rates (ASYRs) allow comparisons of the burden of premature mortality by ethnic group and specific cause of death (Figures 1, 2, 3). For example, for almost every leading cause of premature death in men and women, African Americans had the highest ASYRs compared to other ethnic groups. Among African American men, the disparity in ASYRs was most notable for violent assault (homicide), followed by HIV/AIDS, vascular diseases (ischemic and hypertensive heart, and cerebrovascular disease), accidental drug overdose, and lung cancer. Among African American women, the disparity in ASYRs was most notable for vascular diseases (ischemic and hypertensive heart, and cerebrovascular diseases), breast cancer, HIV/AIDS, and accidental drug overdose.

Discussion
The key findings of this study are that (1) the leading causes of premature mortality were largely preventable: among men, these were HIV/AIDS, suicide, drug overdose, homicide, and alcohol use disorder; and among women, these were lung cancer, breast cancer, hypertensive heart disease, colon cancer, and diabetes mellitus; (2) leading causes of premature death differed remarkably between ethnic groups (Tables A-1–A-4); (3) a large health disparity was measured between African Americans and other ethnic groups: African American age-adjusted overall and cause-specific YLL rates are notably higher, especially for homicide among men (Figures 1, 2, and 3); and (4) except for homicide among Latino men, Latinos and Asians had comparable or lower YLL rates among the leading causes of premature death compared to whites (Figures 2 and 3). These results illustrate how death registry data can be used to measure, rank, and monitor the leading causes of premature mortality for a local geographic region. Such studies can be used to monitor the local mortality burden of disease and injury over time. For example, our results were compared to our previous San Francisco YLL study for the period 1990–1995 [11]. While the burden of HIV/AIDS deaths decreased remarkably, the ethnic health disparities remained, with African Americans continuing to suffer the largest burden. This was especially striking for homicides among African American men. The generally better health status of Asians and Latinos has persisted.

Several of these findings mirror those from national studies [20]. For example, the U.S. Burden of Disease and Injury Study [21] found many of the same preventable causes of premature death among the leading causes, and that the YLL ranking for each ethnic group was unique. Like our study, there were large disparities, measured as DALYs, between African Americans and other ethnic groups, and they reported better health outcomes among...
Asians than whites. The Eight Americas Study [22,23] also found large disparities, measured as life expectancy, between Asian Americans and African Americans. A recent examination of the U.S. black-white disparity in life expectancy during the period 1983–2003 [24] found, like our study, that cardiovascular disease (both males and females), homicide (males), and HIV/AIDS (males) were leading contributors to the gap in recent years.

Three measures were used in this study: YLLs, average YLLs, and ASYRs. The YLL is a stand-alone measure of mortality burden not requiring population estimates. It was used to rank the 15 leading causes of death for men and women (Table 7). However, these 15 leading causes were influenced by the larger number of deaths among older residents. To highlight premature, preventable causes of death, we then ranked these top 15 causes by their average YLLs. Notably, many of the leading causes of death have strong social determinants. Alternatively, the ASYR could have been used to rank the leading causes of death; however, this was not our first choice because it requires population estimates, and the rankings would still be influenced by older deaths. Given our availability of population estimates, ASYRs were used to make comparisons among ethnic groups (Table 6 and Additional file 1). However, only the YLLs (including average YLLs)

By ranking the leading causes of premature death by YLLs, we were able to age-standardize the YLLs to allow leading cause of deaths to be ranked for parts of the population (such as specific ethnicities or geographic areas) for which population estimates are not available. Fourth, subranking by average YLLs identifies leading causes of premature death, bringing attention to preventable deaths that contribute most to the mortality burden. Fifth, these analyses can be repeated periodically to monitor changes, guide and inform policy makers, and to direct and evaluate interventions.

Sixth, except for motor vehicle accidents [16], we used the Global Burden of Disease ICD-10 cause of death categories, making our methods similar to national and international studies [15,21]. Seventh, our study included Latinos/Hispanics, an important segment of the population that was not included in a similar national study [21]. Eighth, with the availability of ethnic-specific population estimates, we were able to age-standardize the YLLs to measure, compare, and monitor the ethnic health disparities in the burden of premature deaths. And ninth, our study findings are directly relevant and can be adapted to the diverse and unique needs of our communities, and to our local government and policymakers.

This study also has several limitations. First, the accuracy of data recorded on death certificates (e.g., underlying cause of death and ethnicity) varies by region and underlying cause [25]. Additionally, analyses using underlying cause of death categories may underestimate the mortality burden for selected contributing causes of death listed on the death certificates (e.g., diabetes mellitus) [26]. Secondly, the YLL metric does not measure well conditions that cause significant disease and disability, but are difficult to measure (e.g., mental illness) or do not result in death (e.g., osteoarthritis). Third, on average, there may be a 10-month or longer delay from the time a calendar year ends and the availability of ICD-10-coded death registry data.

Fourth, the ranking of a specific cause of death depends on its individual YLL magnitude as well as its relative contribution compared to other causes; changes in ranking for a cause over time may be due either to changes in the occurrence of that cause, or to changes in the occurrences of other causes ranked above or below it. Fifth, the average
YLL could be large for a specific cause of death but only involve a small number of deaths (small burden). To avoid this problem, we only evaluated the average YLL for the highest ranked causes of death based on YLLs. Sixth, the YLL measure is not age-standardized and cannot be used to compare specific causes of death between groups with different age compositions. (With population estimates, YLL can be age-standardized as described in Methods.) And seventh, because of the uncertainty of population estimates, age-standardized rates must also be interpreted with caution. In spite of these limitations, using YLLs to rank the leading causes of premature death provides community residents, community-based organizations, policy makers, public health authorities, and researchers with local, representative, objective, and informative data to guide and inform public health priorities, and to direct and evaluate public health interventions.

This study has the following key implications: First, we provide the methodological details for calculating YLL to measure the burden of premature mortality for any geographic area that has death registry data. We provide both the ICD-10 causes of death classifications used for this study [Additional file 2] and the computational program code for calculating age-interval-specific expected years of life lost that can incorporate discounting (used in this study) and age weighting (not used in this study) [Additional file 3]. This code can be executed in a freely available, open source program for statistical computing and graphics [19]. And second, we demonstrate how these results can be used to rank the leading cause of premature death for major ethnic groups. The rankings can be used to guide, inform, and monitor public health priorities and programs for each group. These analyses can become part of routine public health surveillance for local health jurisdictions, as we have done in San Francisco.

Conclusion
Population health measures based on YLLs are readily calculated and useful for measuring, ranking, and monitoring the leading causes of premature death for a local geographic area, and for measuring and monitoring the impact of local efforts to reduce premature mortality in ethnic groups for which there are health disparities.

Competing interests
The author(s) declare that they have no competing interests.

Authors’ contributions
TJA conceived and designed the study, conducted the analyses, and prepared the initial manuscript. DYL reviewed the literature on Global Burden of Disease Study methods and local area research studies, and applied findings to our cause of death classifications. RR assisted in statistical programming, quality control, review of quantitative methods, and development of population health applicability of these measures. MHK reviewed the study for clinical accuracy, epidemiologic methods, and public health impact. All authors contributed substantially to the interpretation of findings and manuscript revisions. All authors read and approved the final manuscript.

Additional material

Additional file 1
Leading causes of premature death by ethnicity and sex, San Francisco, 2003–2004. This file contains tables for the leading causes of premature death by ethnicity and sex. We illustrate how these methods can be repeated for population subgroups to inform and guide public health priorities.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2458-8-116-S1.pdf]

Additional file 2
Cause of death classification using International Classification of Diseases, 10th Revision (ICD-10) codes. This file contains the ICD-10 codes used for our cause of death categories, which were adapted from the World Health Organization Global Burden of Disease Study [15] and the Centers for Disease Control and Prevention External Cause of Injury Mortality Matrix [16].

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2458-8-116-S2.pdf]

Additional file 3
Open source programming code for calculating expected years of life lost. This file contains R programming code to calculate expected years of life lost. R is a comprehensive, open source software package for statistical computing and graphics [19].

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2458-8-116-S3.pdf]

Acknowledgements
Special thanks to Roma Guy (Health Education Department, San Francisco State University; Health Commissioner, City and County of San Francisco) and Virginia Smyly (Deputy Director, Community Programs and Prevention, San Francisco Department of Public Health) for reviewing multiple drafts and providing valuable feedback.

Sources of funding: This study was supported by the San Francisco Department of Public Health (all authors).

References
1. Institute of Medicine, Committee for the Study of the Future of Public Health: The Future of Public Health National Academy Press; 1988.
2. Institute of Medicine, Committee on Assuring the Health of the Public in the 21st Century: Future of the Public’s Health in the 21st Century National Academy Press; 2003.
3. Centers for Disease Control and Prevention: Updated guidelines for evaluating public health surveillance systems: Recommendations from the guidelines working group. MMWR Recomm Rep 2001, 50(RR-13):1-51.

4. Murray CJL, Lopez AD, (Eds): A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020 (The Global Burden of Disease and Injury) Harvard School of Public Health; 1996.

5. Kominski GF, Simon PA, Ho A, Luck J, Lim YW, Fielding JE: Assessing the burden of disease and injury in Los Angeles County using disability-adjusted life years. Public Health Rep 2002, 117(2):185-191.

6. Marshall RJ: Standard expected years of life lost as a measure of mortality: norms and reference to New Zealand data. Aust N Z J Public Health 2004, 28(5):452-457.

7. Mariotti S, D’Errigo P, Mastroeni S, Freeman K: Years of life lost due to premature mortality in Italy. Eur J Epidemiol 2003, 18(5):513-521.

8. Katzenellenbogen J, S SPS: Western Australian Burden of Disease Study: Mortality 2000. Tech Rep 2003 [http://www.health.wa.gov.au/publications/documents/WA_Burden_of%Disease.pdf]. Department of Health, Perth, Western Australia.

9. Centers for Disease Control and Prevention: Years of potential life lost before age 65-United States, 1990 and 1991. MMWR Morb Mortal Wkly Rep 1993, 42(13):251-253.

10. Gardner JW, Sanborn JS: Years of potential life lost (YPLL)-What does it measure? Epidemiology 1990, 1(4):322-329.

11. Aragon T, Reiter R, Katcher B: San Francisco Burden of Disease and Injury: Mortality Analysis, 1990–1995. Tech Rep 1998 [http://www.sfdph.org/phd/files/reports/Studies/Data/DiseaseInjury/bd99095b1.pdf]. San Francisco Department of Public Health.

12. Centers for Disease Control and Prevention, Center for Health Statistics: [http://www.dhs.ca.gov/hisp/chs/default.htm]. Department of Health, Perth, Western Australia.

13. State of California, Department of Finance: San Francisco Race/Ethnic Projections with Age and Sex Detail, 2000 – 2050. [http://www.dof.ca.gov/HTML-DEMOGRAPHIC/Data/RaceEthnic/Population-00-50/documents/SanFrancisco.txt].

14. World Health Organization: International Classification of Diseases (ICD), 10th Revision. 2007 [http://www.who.int/classifications/apps/icd/icd10online/].

15. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL, (Eds): Global Burden of Disease and Risk Factors 1st edition. 2006 [http://www.dcp2.org/pubs/GBD]. World Bank Publications.

16. Centers for Disease Control and Prevention, Center for Health Statistics: ICD Framework: External Cause of Injury Mortality Matrix. 2002.

17. van Dommelen L: Linear interpolation. 2007 [http://www.epf.fsu.edu/~dommelen/courses/eml3100/aids/intpol/index.html]. [Accessed May 23, 2007].

18. Anderson RN, Rosenberg HM: Age standardization of death rates: implementation of the year 2000 standard. Natl Vital Stat Rep 1998, 47(3):1-16.

19. R Development Team: R – A language and environment for statistical computing and graphics. [http://www.r-project.org/].

20. WONG MD, Shapiro HF, Boscardin WJ, Ettenot SL: Contribution of major diseases to disparities in mortality. N Engl J Med 2002, 347(20):1585-1592. [http://dx.doi.org/10.1056/NEJMsa012979].

21. Michaud CM, McKenna MT, Beggs S, Tomijima N, Majmudar M, Bulzacchelli MT, Ebrahim S, Ezzati M, Salomon JA, Kreiser JG, Hogan M, Murray CJ: The burden of disease and injury in the United States 1990–1996. Popul Health Metr 2006, 4:11.

22. Murray CJL, Kulkarni S, Ezzati M: Eight Americas: new perspectives on U.S. health disparities. Am J Prev Med 2005, 29(5 Suppl 1):4-10.

23. Murray CJL, Kulkarni SC, Michaud C, Tomijima N, Bulzacchelli MT, Iandorio TJ, Ezzati M: Eight Americas: investigating mortality disparities across races, counties, and race-counties in the United States. PLoS Med 2006, 3(9):e260.

24. Harper S, Lynch J, Burris S, Smith GD: Trends in the black-white life expectancy gap in the United States, 1983–2003. JAMA 2007, 297(11):1224-1232.

25. Maudsley G, Williams EM: “Inaccuracy” in death certification: are we now? J Public Health Med 1996, 18:59-66.

26. Redelings MD, Sorvillo F, Simon P: A comparison of underlying cause and multiple causes of death: US vital statistics, 2000–2001. Epidemiology 2006, 17:100-103.

**Pre-publication history**

The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1471-2458/8/116/prepub