Cross Cutting
Health Disparities

GUIDE FOR SUB-COUNTY ASSESSMENT OF LIFE EXPECTANCY (SCALE)

Improving the Capacity of States and Local Health Departments to Understand Life Expectancy Disparities

CSTE
COUNCIL OF STATE AND TERRITORIAL EPIDEMIOLOGISTS

Version 2.1
September 2017

Table of Contents

1.0. Executive Summary

As communities become increasingly data savvy, residents want to know what health is like in their area or neighborhood. Looking at data at a national, state, or even county-level data hides a wide range of values, including some stark disparities. However, many local health jurisdictions have lacked the capacity to examine sub-county data. Recent results show examining neighborhood level estimates of Life Expectancy (LE) at birth in the context of known behavioral, social, and environmental risk and protective factors has been effective at reaching the community and gaining resources to address areas of concern. This **Sub-County Assessment of Life Expectancy** (SCALE) Guide is intended to serve as a resource for public health practitioners and their partners who are working to identify, measure, and understand growing and persistent community level health disparities and to catalyze collective actions to address the underlying causes of these disparities. The SCALE project goal is to improve the capacity of states and large local health departments to calculate sub-county level LE estimates.

LE at birth is defined as the estimated number of years a newborn can expect to live if current age-specific death rates in that population remained the same over time [1]. This measure is particularly useful for examining community-level disparities because it reflects the impact of major illnesses and injuries and their underlying causes, enables direct comparisons across geographies and time, and is simpler and more intuitive to the public and policy makers than are other measures of death (e.g., standardized mortality ratios, age-adjusted mortality rates, and years of potential life lost) [2 -7].

Scaling these efforts across the U.S. can inform future research and focus attention of policy makers, legislators, and the public on underlying conditions that are immediately actionable. To advance this initiative, the Council of State and Territorial Epidemiologists (CSTE), Centers for Disease Control and Prevention (CDC), six state (Florida, Massachusetts, Maine, New York, Washington, and Wisconsin) and two local (Los Angeles County and Public Health, Seattle & King County) health departments reviewed existing literature and methods, identified software tools, and developed this draft Guide.

The Guide is arranged to provide the background rationale for sub-county methods used in SCALE; explain how, what and where to find data for LE calculations; share an existing tool; and show how State and Local Health Departments utilized the process and what outcomes were experienced.

Calculation of LE can enable the following future public health practice and research applications:

1. Identify and monitor community hot spots of health disparities.
2. Visually examine the degree to which LE and associated contributing factors vary across populations and geographies.
3. Raise public awareness about the importance of place-based factors in creating health and health disparities including those not traditionally associated with public health (i.e., education, housing, transportation, community development, and employment).
4. Facilitate research on the relative contribution of specific behavioral, social, and environmental factors in creating health.
5. Catalyze multisector collaborations and empower communities to more effectively address upstream factors, reduce disparities, and improve community health.

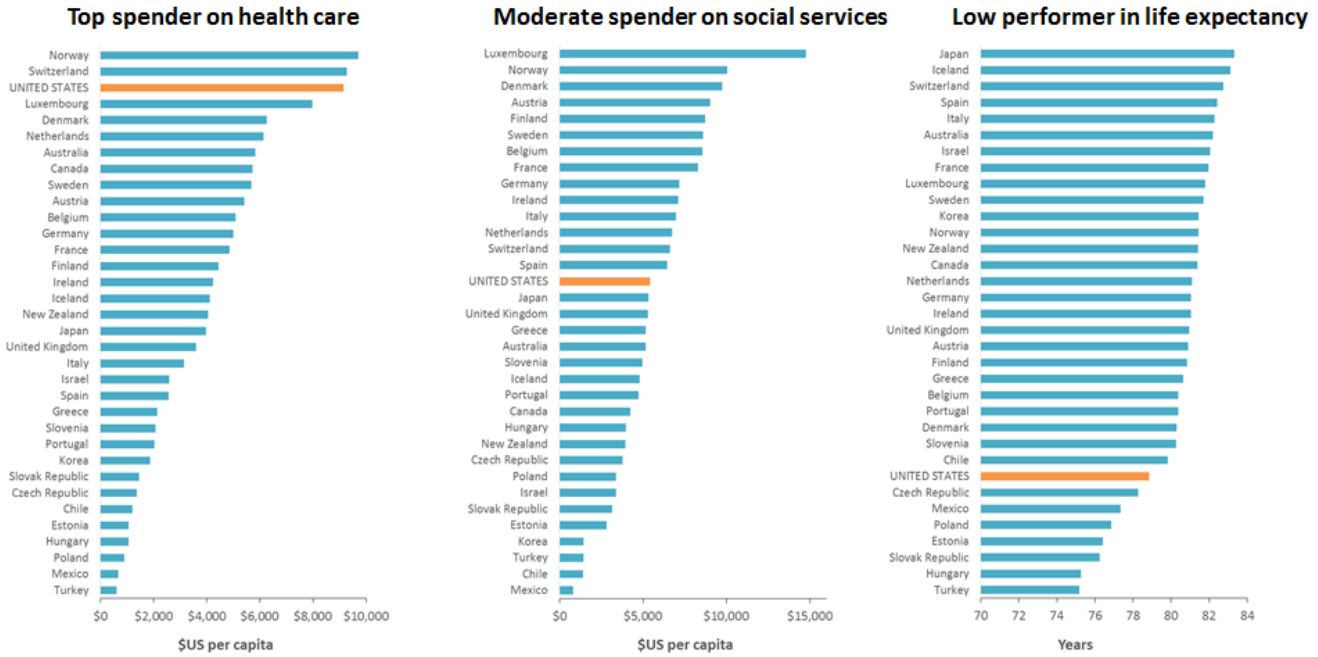
2.0. Introduction

LE is a measure that holds many people's attention, and is a concept that is easy for community to grasp. LE at birth is defined as the estimated number of years a newborn can expect to live if current age-specific death rates in that population remained the same over time [1], and is particularly useful for examining community-level disparities because it reflects the impact of major illnesses and injuries and their underlying causes, enables direct comparisons across geographies and time, and is simpler and more intuitive to the public and policy makers than are other measures of death (e.g., standardized mortality ratios, age-adjusted mortality rates, and years of potential life lost) [2 -7]. All states in the U.S. are required to routinely and systematically report death, and information from the death certificates (race/ethnicity, age, and a geographic identifier such as address, city, or ZIP code) can readily be used to calculate reliable and comparable LE estimates. Communities can galvanize around seeing a gap of 15 or 20 years between the longest lived and the shortest lived areas. In many cases, local public health may already have created calculations of LE at the county or state level, as those are often straightforward and may not require a detailed look at the underlying data going into the calculation. When looking at smaller geographies, however, there are a number of issues to keep in mind, particularly as numbers get small.

2.1. Disparities in Life Expectancy at the Country, County, and Local Levels

Disparities in LE estimates between the U.S. and other countries in the context of health expenditures have attracted increasing attention during the past few years. In 2010, the U.S. ranked 40th for male and 39th for female life expectancy at birth among 187 countries [8] [9], even though the U.S. spends almost twice as much per capita on health care than does any other country (Figure 1) [9][10]. All Americans, even the most educated, affluent, and well-insured, live sicker lives than those in other developed countries [11 -14]. More disturbing, the gap appears to be widening. Comparisons of historical trends of life expectancy between the U.S. and other countries found that, since ranking seventh in life expectancy during the 1950s, the U.S. has dropped more than 25 places, with the most rapid relative declines occurring during the past three decades among women [12] [15]. According to a 2012 Annual Review of Public Health article, this lag in U.S. health status results from "structural factors related to inequality and conditions of early life" [9].

Figure 1: Health Care and Social Services Spending versus Life Expectancy, by Country



Sources: Health expenditures per capita, 2013 and life expectancy at birth, 2013 (World Bank); Social expenditures per capita, 2011 (OECD)

[1]

Results of multiple studies suggest the primary driver of the poor relative national level performance of the U.S. on several public health indicators, including those seen in Figure 1, is profound disparities in LE at birth across U.S. counties [15-17]. For example, LE in some top performing counties— females in Marin County, California (85.0 years) and Montgomery County, Maryland (84.9 years) and males in Fairfax County, Virginia (81.7 years), and Gunnison County, Colorado (81.7 years)—is comparable with LE in countries where populations live the longest including Japan and Switzerland. In contrast, LE estimates for males in McDowell County, West Virginia (63.9 years), and Bolivar County, Mississippi (65.0), and for females in Perry County, Kentucky (72.7), and Tunica County, Mississippi (73.4), were lower than estimates for Algeria, Bangladesh, and Nicaragua [18].

The U.S. LE remains significantly lower than the longest lived countries. In 2009, the U.S LE was 78.2 years, compared to 81.8 in the longest lived countries. The current U.S. LE is equivalent to the LE that the 10 longest lived countries had in 1990. Extrapolating from this finding, it would take 19 years of improvement for the U.S. to catch up to the current LE in the longest lived countries. Similar comparisons demonstrate that disparities among counties when compared to the longest lived countries were even greater than between nations (Figure 3). LE ranged from 16 years *ahead* of the longest lived countries to more than 50 years behind [18]. If current trends hold, the worst performing counties don't have much of a chance to catch up to the best performing countries OR counties, as the top performing counties have seen steady gains over time, whereas LE in the worst performing counties have stagnated over the past 25 years [19].

[1] http://data.worldbank.org/, accessed July 2015

Figure 2: Historic and Projected Life Expectancy of the Longest-lived Countries, by Year, 1950 to 2050[2]

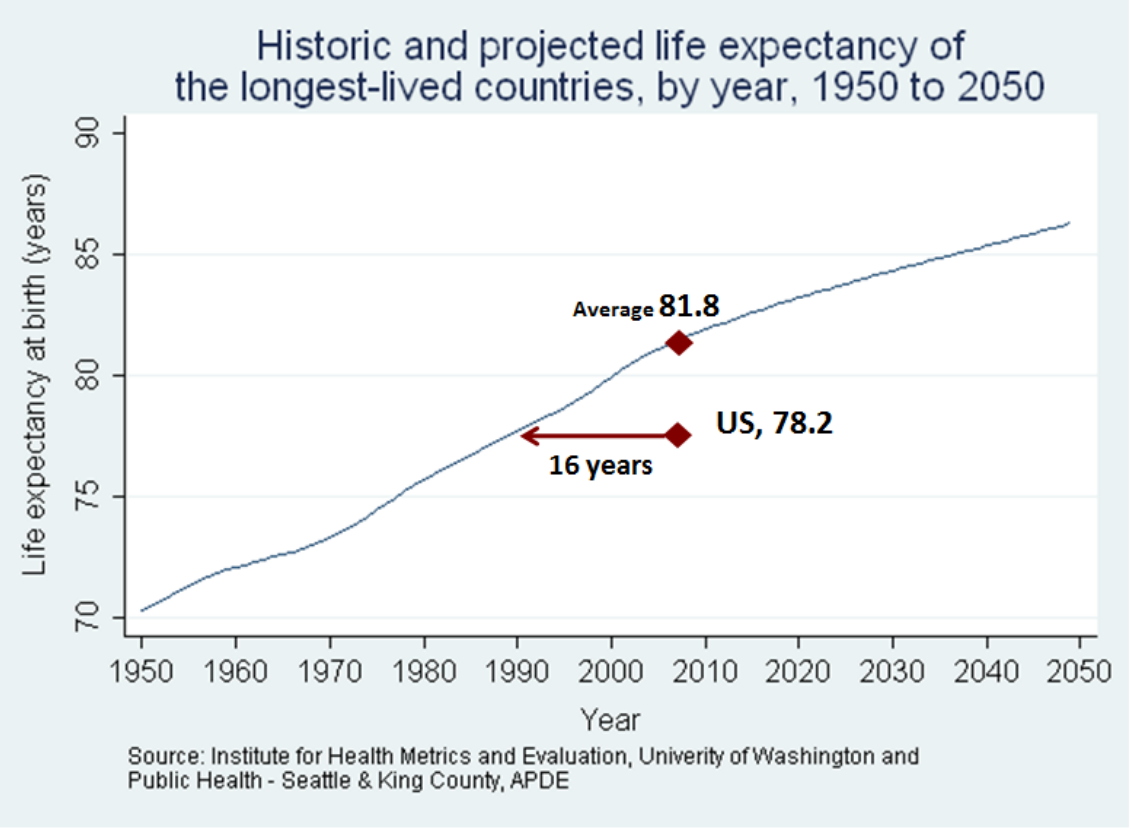
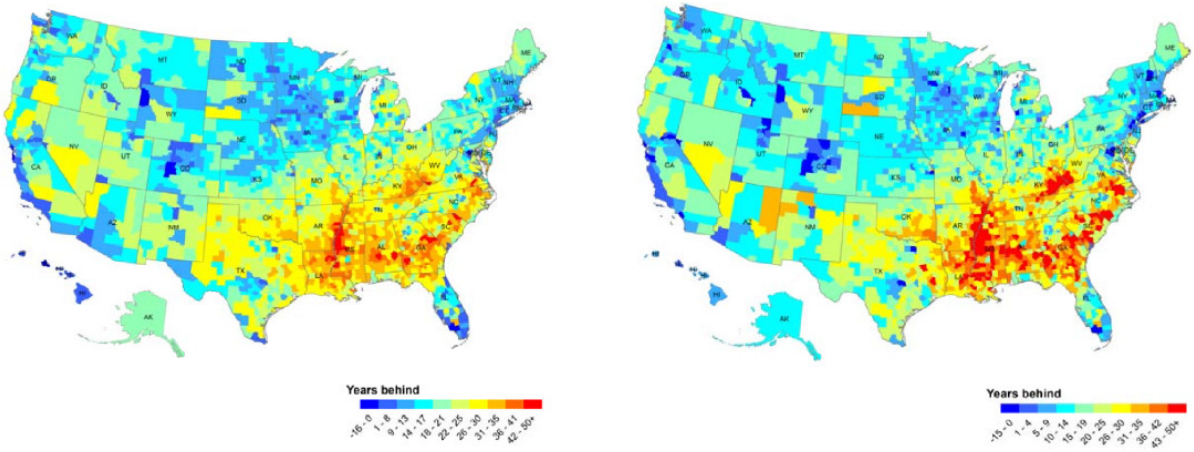


Figure 3: Life Expectancy by U.S. County and by the 10 Countries with the Highest Life Expectancy[18]



Females Males

As seen in the comparisons above, the national estimates mask county-level disparities in LE. Recent evidence suggests that county-level LE measures are masking similar magnitudes of disparities at the

[2] Presentation to King County Board of Health, Assessment, Policy Development & Evaluation, 5/2013

sub-county level, even in counties that perform well overall on other measures of premature death. For example, researchers reported that the incidence of premature death in Boston was 1.39 times higher (95% CI 1.09–1.78) for persons living in census tracts where >20% of the population had incomes below the federal poverty level than it was for census tracts where <5% of the population lived in poverty. Similarly, the results of a study examining health disparities in 77 communities within Chicago found LE estimates varied by more than 15 years, ranging from 68.2 to 83.3 years [20]. As demonstrated in these examples, considering the geographic context of premature death has enormous potential for identifying local concentrated areas (or "hot spots") of health disparities and facilitating research on the role of local area factors including housing, education, employment opportunities, environmental conditions, behavioral factors, and access to health care and material goods that impact social disparities in health [21].

> RIGHT HERE IN KING COUNTY WE HAVE SOME OF THE BIGGEST HEALTH DISPARITIES YOU'LL FIND ANYWHERE. PEOPLE WHO LIVE IN CERTAIN KING COUNTY NEIGHBORHOODS ENJOY SOME OF THE LONGEST AND HEALTHIEST LIVES OF PEOPLE ANYWHERE, WHILE JUST A BIKE RIDE AWAY LIFE EXPECTANCY AND QUALITY OF LIFE IS MUCH MORE SIMILAR TO DEVELOPING COUNTRIES. THERE IS PLENTY THAT CAN BE DONE, AND MANY WAYS TO BE PART OF THE SOLUTION BOTH LOCALLY AND GLOBALLY.
>
> David Fleming, Former Director of Public Health Seattle King County

2.2. Interest in and Need for Life Expectancy Estimates

These large magnitudes of community-level disparities have caught the attention of U.S. legislators. For example, Senator Bernie Sanders, who chaired the Senate Subcommittee on Primary Health and Aging, held a congressional hearing in 2013 titled, "Dying Young: Why Your Social and Economic Status May Be a Death Sentence in America." The hearing included testimony from physician and research experts on health, economic, and educational factors that contribute to disparities in LE. At the hearing's conclusion, Senator Sanders cited poorer parts of the U.S., including some rural counties and inner-city neighborhoods, noting, "In many ways, the stress of poverty is a death sentence, which results in significantly shorter life expectancy. Parts of Boston and Baltimore have a lower life expectancy than Ethiopia and Sudan."

In addition to the increased legislator and public attention to community-level health disparities, several recent developments have increased the demand for assessing and improving local population health. First, the voluntary public health accreditation standards[3], launched in 2011, require a comprehensive community health assessment and community improvement plan as prerequisites for state and local health departments seeking accreditation. Second, Section 9007 of the Affordable Care Act requires that the >3,000 nonprofit hospitals across the U.S. complete a community health needs assessment every 3 years and adopt an implementation strategy to meet identified needs in order to

[3] http://www.phaboard.org/

continue to meet tax-exemption status. One specific requirement of the accreditation standards and the Internal Revenue Service regulations is identification of and engagement with community members or their representatives from populations experiencing health disparities within their jurisdictions. Finally, Public Health 3.0 [4] challenges local health to serve as a chief health strategist, looking at social determinants of health, and examining data at a local level, with community context.

3.0. Examples from the Field, Prior to SCALE

Several health departments have successfully used sub-county estimates of LE at birth to identify and explore local hot spots of health disparities, to raise public awareness, and to catalyze multisector partnerships and collective actions. In Sections 3.1 and 3.2, we present case studies from two such health departments, Public Health—Seattle & King County and the Los Angeles County Department of Public Health to provide examples of the promise and utility of sub-county life expectancy estimates for informing public health action. In addition, the Robert Wood Johnson Foundation (RWJF) has produced a series of story maps to catalyze conversations about the inequity in LE. We provide a brief summary of the RWJF effort in Section 3.3

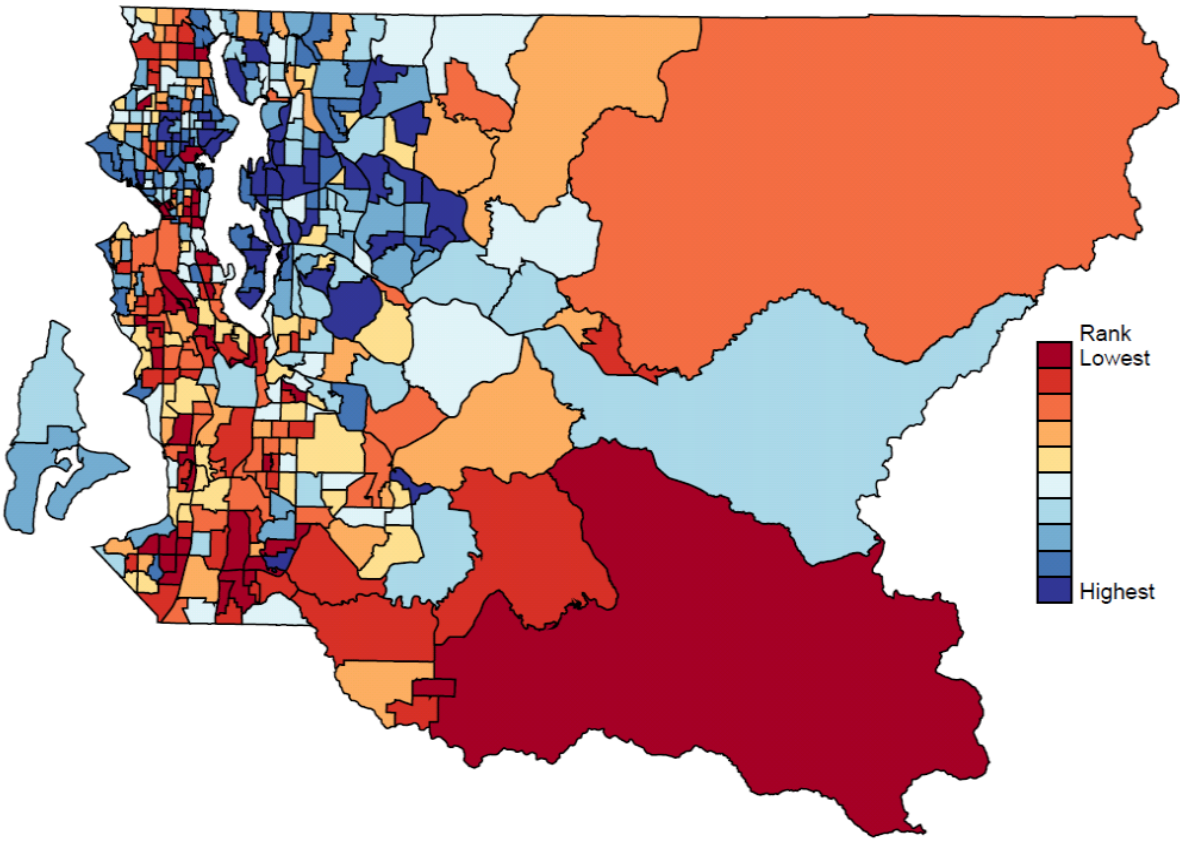
3.1. Public Health, Seattle & King County

King County, home to 2.1 million residents and 39 cities, is the most populous county in Washington State. Home to businesses such as Microsoft, Amazon, Weyhauser, and sporting other technology hubs, King County health outcomes and risk behaviors tend to compare favorably to other counties in the US; however, health equity work has shown the high performing county rate masks large disparities in health [27]. In 2012, Public Health, Seattle & King County (PHSKC) calculated LE at a census tract level for the 398 tracts in King County to begin examining place-based disparities. The 5-year estimates showed a range of 25 years (after suppression of unreliable rates), with a low of 72 years and a high of 96 years. The overall King County LE was 81.6 years.

These findings generated questions from local leaders and community members how additional health behaviors and health outcomes would look at a similar geography, and PHSKC embarked on a small area estimation project that showed a consistent pattern of disparities.[23] This work was presented in 2013 at a Federal Reserve Bank meeting, culminating in a place-based initiative called Communities of Opportunity (COO),[5] a public-private partnership with the Seattle Foundation and Living Cities, and is a cross-divisional initiative with PHSKC and the Department of Community and Human Services in King County. COO's goal is to create greater health, social, economic and racial equity in King County so that all people thrive and prosper, regardless of race or place, with a focus on economic, health, housing, and community metrics. The project is rooted in the community, using a collective impact framework that allows the community to shape their own solutions. The COO project is also aligned within the King County Accountable Community of Health.[6]

[4] DeSalvo, K. Public Health 3.0: Time for an Upgrade., AJPH 106(4), pp. 621–622
[5] http://www.kingcounty.gov/elected/executive/health-human-services-transformation/coo.aspx, last checked 2/2017
[6] http://www.kingcounty.gov/elected/executive/health-human-services-transformation/ach.aspx

Figure 4. Life Expectancy at Birth, King County Census Tracts, 2008-2012

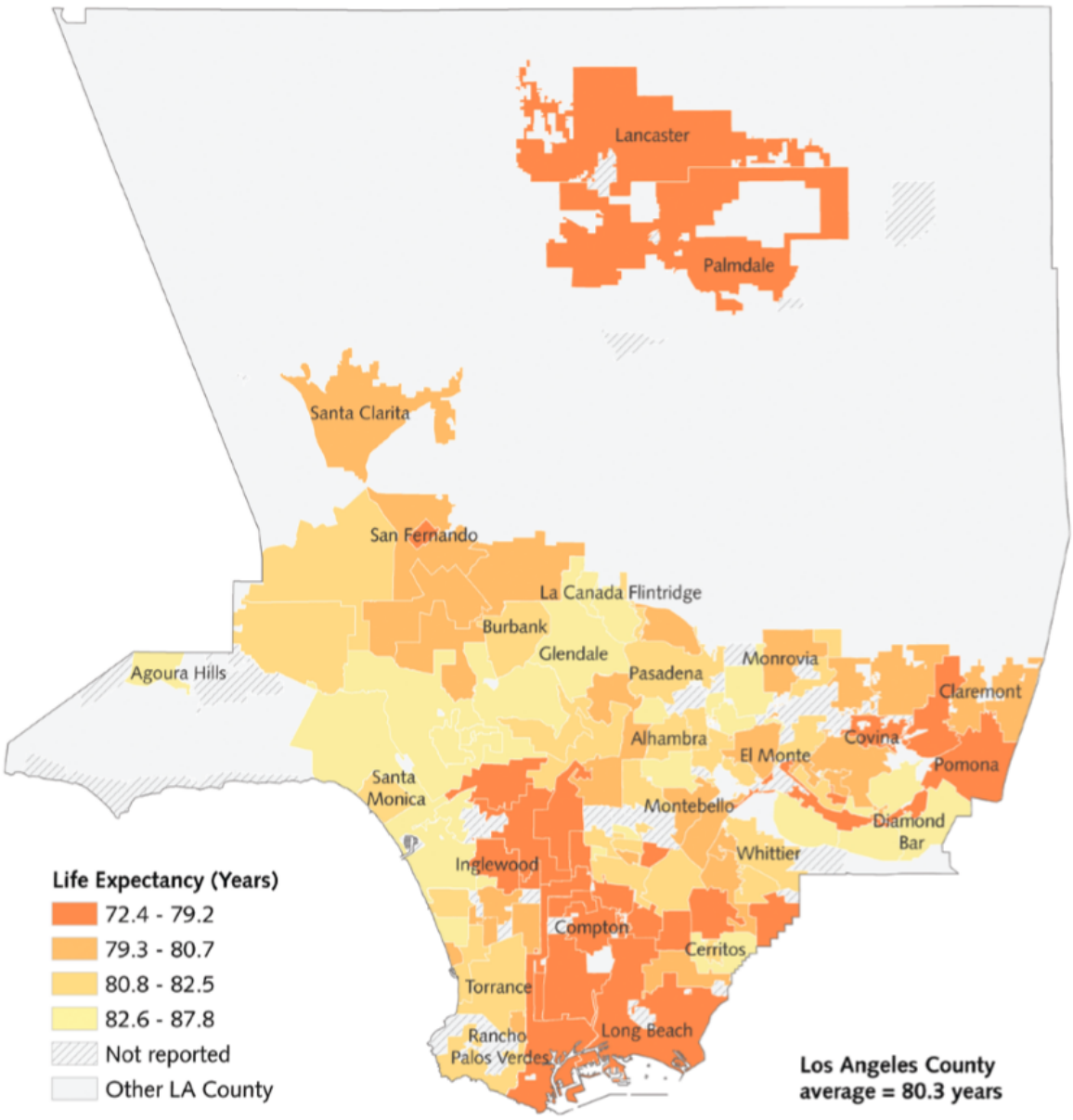


3.2. Los Angeles County

In 2009, the Los Angeles County Department of Public Health (LACDPH) calculated LE at birth for 103 cities and communities within the County.[2] In earlier analyses, large and persistent disparities in LE had been observed, and the LACDPH recognized that there was a need to bring increased attention to addressing the underlying social and physical determinants of health in order to make progress in narrowing these disparities. The County's cities and unincorporated communities were viewed as important partners in this effort, and it was hoped that examining LE at the city and community level would bring increased attention and engagement.

LE across cities and communities varied widely, ranging from 72.4 years to 87.6 years, and was strongly correlated with community-level economic hardship. Cities/communities were ranked by LE and by economic hardship, and this information was published in a report that was published and broadly disseminated to the general public, city mayors, councilmembers, city planners, and as other public health stakeholders. The information resulted in increased engagement with communities, local governments, policymakers, city planners, and other sectors; it also increased recognition of the important impacts of the physical and social environments on health.

Figure 5. Life Expectancy at Birth in LA County Neighborhoods



3.3 Robert Wood Johnson Foundation LE work

Similarly, RWJF has generated public attention on community-level health disparities by funding Virginia Commonwealth University to create a series of maps showing LE in 20 U.S. cities. These maps depict dramatic differences in LE in several jurisdictions. For example, the project demonstrated disparities by as much as 25 years within New Orleans (Figure 6).

Figure 6: Metro Map: New Orleans, Louisiana.[7] *The average LE for babies born to mothers in New Orleans varies by as much as 25 years just a few miles apart.*

4.0. SCALE Project

To address the needs articulated in Section 2.0 and to scale the successes demonstrated by, PHSKC, Los Angeles County, and other jurisdictions, the Centers for Disease Control and Prevention (CDC) provided funds to the Council of State and Territorial Epidemiologists (CSTE) in October 2014 to engage several state and local health departments in a multi-year project. This project, now entitled the Sub-County Assessment of Life Expectancy (SCALE) Project, has engaged several jurisdictions to date in developing and pilot testing resources to support state and local health departments throughout the U.S. in calculating sub-county LE estimates. In this section, we describe the goals and key activities that comprise the SCALE Project.

4.1. SCALE Project Goal

The SCALE project goal is to improve the capacity of states and local health departments to calculate sub-county–level LE at birth estimates. Calculation of LE estimates in many jurisdictions throughout the U.S. can enable the following future public health practice and research applications:

1. Identify and monitor community hot spots of health disparities.
2. Visually examine the degree to which LE and associated contributing factors vary across populations and geographic locations.

[7] http://www.rwjf.org/en/library/articles-and-news/2015/09/city-maps.html

3. Raise public awareness about the importance of place-based factors in creating health and health disparities including those not traditionally associated with public health (i.e., education, housing, transportation, community development, and employment)
4. Facilitate research on the relative contributions of specific behavioral, social, and environmental factors to life expectancy.
5. Catalyze multisector collaborations and empowered communities to more effectively address upstream determinants of health, reduce disparities, and improve community health.

4.2. SCALE Project Activities

The SCALE project aims to improve the capacity of state and local health departments to calculate sub-county LE estimates by encouraging participation in the project, developing and disseminating easy-to-use resources, and identifying and sharing lessons learned from the project through evaluation activities (See Table 1). SCALE was created as a multi-phase effort in which the work of each phase builds upon the last. The project commenced in January 2015 when CSTE and CDC invited six state health departments (Florida, Massachusetts, Maine, New York, Washington, and Wisconsin) and two local health departments (LACDPH and PHSKC) with previous experience in small area analysis to participate in the first phase of the project.

During Phase I, the eight jurisdictions engaged in collaborative efforts to identify via literature review, test, and suggest methods for calculating sub-county LE, and produced a guidance document (this "Guide") to share lessons learned and decisions point guidance with other jurisdictions. These eight jurisdictions are the core Workgroup, which meets on a regular basis to discuss current updates to LE work, additional methods, and to provide TA to incoming jurisdictions.

In Phase II, 25 additional state and local health departments of varying sizes joined the SCALE project (Figure 7). These jurisdictions were tasked with pilot testing the draft Guide produced during Phase I and assessing the extent to which the methods and associated tool identified in Phase I (i.e., the SEPHO tool, see Section 5.0) met their needs in calculating sub-county LE estimates.

In addition, states and localities from Phase I and a subset of Phase II participants engaged in discussions and activities to address several methodological questions raised during Phase I, such as how to address cells with zero deaths and how to treat areas where a large proportion of the population lives in group quarters. Initial efforts to identify best practices for mapping sub-county LE estimates drawing upon expertise from CDC's Geospatial Research, Analysis, and Services Program (GRASP) were also undertaken during Phase II.

Additional states and localities will join Phase III of the effort in Fall 2017 to continue testing and refining methodologies identified or developed in Phases I and II. In addition, during this final phase specific efforts will continue to develop recommendations for visualizing sub-county LE and for communicating about these estimates with various target audiences. Lessons learned from Phase III and recommendations from these efforts will be included in future versions of this Guide.

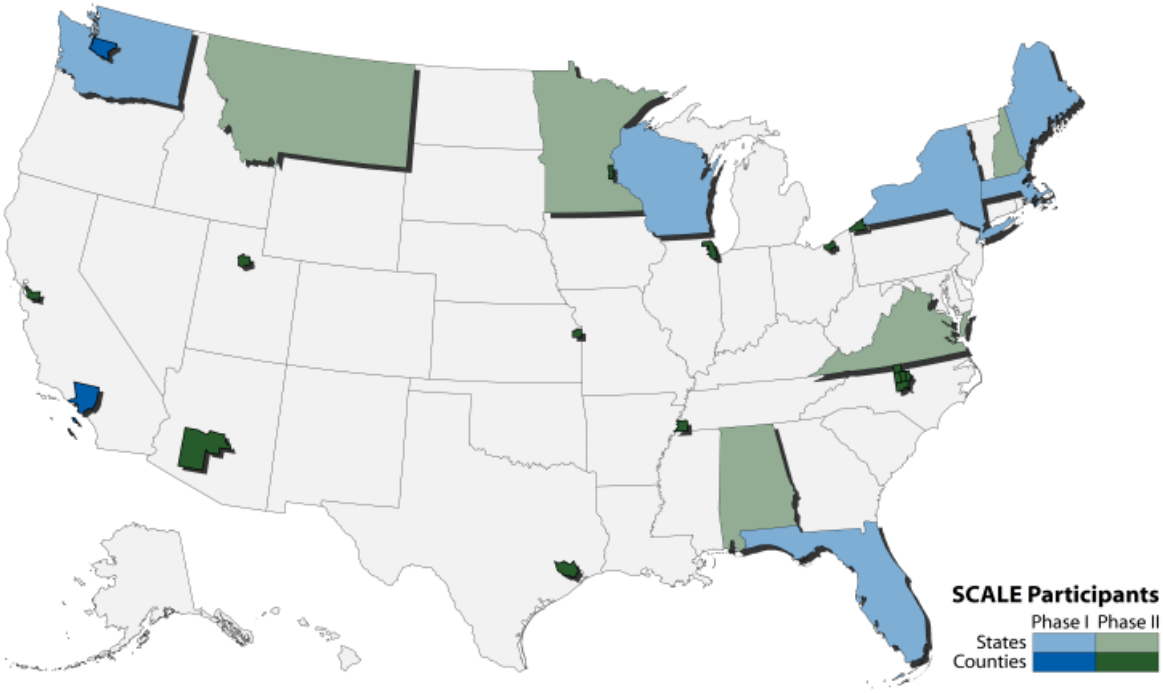
Table 1. SCALE: Project Phases[8]

Phase I	Conducted a literature review to understand the approaches, available parameters, and lessons learned from previous efforts associated with constructing small-area LE estimates.
	Reviewed common approaches used in the literature for calculating direct small-area LE estimates and arriving at initial decisions about methodology.
	Identified other existing tools for calculating LE that might easily be adopted/adapted (SEPHO).
	Compared calculations produced by SEPHO tool with other methodologies for generating LE estimates (SAS and STATA code from previous LE efforts), refined approach.
	Developed evaluation plan for Phase II.
	Products include: (1) Draft Guide for state/local health departments with SEPHO tool as approach used, (2) Sub-county estimates for Phase I states/localities, (3) 2015 CSTE conference presentation, (4) Evaluation plan
Phase II	Recruit and orient new states/localities to methodology and general project purpose/approach.
	Implement evaluation; New state/localities pilot test draft materials from Phase I and provide feedback through the evaluation.
	States/localities assess potential refinements in methodology to expand geographic coverage by performing several sensitivity analyses.
	In collaboration with the Geospatial Research, Analysis, and Services Program (GRASP) the Agency for Toxic Substances and Disease Registry (ATSDR), identify methods for visualizing LE using direct estimates of LE.
	Engage expert panel to develop initial recommendations for visualizing and communicating about LE estimates.
	Anticipated products include: (1) Revised tools for estimating LE, (2) Revised/updated Guide, (3) Recommendations regarding visualization and messaging, (4) 2016 CSTE conference presentations, (5) Evaluation findings, (6) Manuscript(s)
Phase III	Recruit and orient new states/localities to methodology and general project purpose/approach.
	Anticipated products include: (1) Revised/updated Guide, (2) Additional tools for LE calculation or consideration, (3) 20167 CSTE conference presentations, (4) Evaluation findings

LE: Life expectancy, SEPHO: South East Public Health Observatory, CSTE: Council of State and Territorial Epidemiologists, SCALE: Sub-County Assessment of Life Expectancy

[8] Paper, in progress

Figure 7. States and Local Health Departments participating in SCALE



5.0. Review of Approaches for Calculating Life Expectancy

Definition of Life Expectancy

Life expectancy (LE) is a summary mortality measure often used to describe the overall health status of a population. For any given population, LE can be calculated at any age (e.g., birth, age 50 years, age 65 years). The SCALE project focuses on LE at birth, which is defined as the estimated number of years a newborn can expect to live if current age-specific death rates in that population remained the same over time [1].

In the U.S., LE is a commonly used indicator of population health and health disparities. Because all states require deaths to be routinely and systematically reported, information from the death certificates (race/ethnicity, age, and a geographic identifier such as address, city, or ZIP code) can readily be used to calculate reliable and comparable LE estimates.

5.1 Overview of approaches considered to develop the life expectancy calculation

Multiple methods exist for estimating LE. These include methods based on stable population concepts, biological theories of aging, estimation of population by age, regression equation methods that exploit the relationship between LE and other demographic indices, construction of abridged life tables and methods that combine traditional complete life table construction techniques with smoothing or graduation methods [22] [24]. After literature review and group discussion, the SCALE Workgroup decided to use an abridged life table method, and the Chiang II calculations. Details are provided below.

Life Tables

A life table shows the probabilities of a member of a particular age dying before their next birthday. In the U.S., two types of life tables are used: the cohort (or generation) life table and the period (or current) life table. The cohort life table is based on age-specific death rates observed through consecutive calendar years and reflects the mortality experience of an *actual* cohort from birth until no one from the group is alive [26]. The period life table represents the mortality experience of a *hypothetical* birth cohort if it experienced throughout its entire life the mortality conditions of the period of interest. The period life table can be considered "a snapshot of current mortality experience and shows the long-range implications of a set of age-specific death rates that prevailed in a given year" [26].

CDC's National Center for Health Statistics publishes complete period life tables annually.[9] Given the routine nature and availability of a nationally published period life table, the Workgroup settled on using this as the basis for Phase I of the SCALE project.

Abridged Life Table

A *complete life table* contains data for every year of age, whereas an *abridged life table* typically contains data by 5- or 10-year age intervals. The SCALE project suggests using an abridged life table with 5-year age intervals except for the first interval, which is set at 0–1 year, and the last interval, which is defined as 85+ years. An abridged table is recommended for smaller areas in recognition of the fact that sub-county geographies would have too many zeros using single age categories. It's important to separate the infant deaths from the 1-4 age category, as infants who die have a much higher rate of death in the first 28 days, compared to other ages, which have a more normal distribution of deaths across the year and age-group. Grouping infant deaths with the 1-4 category will result in a higher LE estimate. , The abridged life table method can be used for any geographic area, including census tracts, ZIP codes, city boundaries, or other geopolitical units.

Adjusted Chiang II Methods

The long-established Chiang method for estimating LE by using a period (current) life table has been widely used internationally [7] [27]. The Chiang method and its variations assume that deaths are spread evenly throughout each age period, except for persons <1 year of age, for whom deaths are highly skewed toward the first 28 days of life. For all other age groups, Chiang assumes a 0.5 age interval; the <1 group is valued at 0.1. One major concern about the Chiang method is age groups for which no deaths occur, which causes a miscalculation in standard error. Therefore, researchers have developed alternative methods to address this issue, including the Chiang II method and the adjusted Chiang II method [10]. One way to prevent zero deaths in a single year of age is to collapse age categories, which is why the abridged life table is suggested.

The adjusted Chiang II method was proposed to modify the assumption in the Chiang I method of a zero variance for the final age band. The adjusted Chiang II method uses the formula by Silcocks to adjust for variance in the final age band [3].

---

[9] http://www.cdc.gov/nchs/products/life_tables.htm.

5.2 Addressing small-area methodologic issues

Calculation of LE at the county or state level is typically straightforward, given generally large population numerators and denominators, and may not require a detailed look at the underlying data going into the calculation. County and states are geographies routinely assigned, and there are often population estimates available at a state or county level. When looking at smaller geographies, however, there are a number of issues that must be considered, as small numbers (deaths or population) or unequally distributed data (deaths or population) may influence the LE calculations.

Small Populations/Minimum Population Size

Several authors have examined the impact of small populations on life expectancy. Variations on the suggested minimum population size range from 3,750 to 7,000. Additional information can be found in Appendix A. Several researchers have suggested that life table estimates overestimate LE for populations less than 5,000 years of life at risk [1] [28] [29]. For the SCALE Phase I project, several Workgroup members assessed the minimum population size for SCALE applications by generating estimates for all tracts, and examined standard error and special conditions of the tract to assess whether a unique feature of the tract (e.g., nursing home residents, incarcerated populations, university students) affected the LE estimate. Others used an R-based tool to aggregate tracts to the minimum population size. One Workgroup member excluded tracts that had a high percentage (>=50%) living in group quarters, tracts that had no land area (were only water), and military bases. In one Workgroup case, tracts remained too small and they used Minor Civil Divisions (MCDs).

Because death data rarely come with assigned geocodes, jurisdictions may need to use address information to geocode the death to tract or other geography (see Section 6, below). Once the geographic assignment is made, numerators and denominators can be evaluated to see whether aggregation of tracts or another higher level geography is helpful to present stable rates.

Standard error and confidence intervals

Because LE can be a tightly grouped dataset, determining a meaningful difference between LE ranges is important. Most of the reviewed papers did not discuss a specific standard error. The SCALE Phase I Workgroup recommends using a standard error of ±2, based on literature review and after analysis of calculated LE for local jurisdictions. The Workgroup suggests considering suppressing LEs with a large standard error.

Zero Cells

Based on the population, population distribution, and death rates, zero cells can occur, even when grouping ages and across years, especially when small geographic units, such as census tract, are used. A zero death count gives an estimate of zero for an age interval, which causes underestimation of the variation; the more zeros and the more underestimation that occur, the larger the underestimation of standard error [25]. Several corrections have been suggested to address the concern, including small substitutions of values for zero and expected numbers of deaths [1] [7] [26] [28]. The biggest concern is when there is a zero count in the <1 or the oldest age band, as those have the largest impact on rates. At this time, SCALE Workgroup view each of the options as acceptable, depending on the jurisdiction's choices. New York State chose to use an R-based tool to group tracts to hit a threshold of 60 deaths to avoid zero cells.

Age 85+ Year Category

A death count of zero in the 85+ year age category would strongly affect LE because the Chiang calculation would give the cohort an infinite survival, raising the LE estimate and standard error. To address this, some options include: replacing the zero cells with a national death rate for the country or a national age-specific death rate for the country [7] [28], or increasing the last age band to 75+. Most Workgroup participants did not have issues with zero cells in the 85+ age category given the 5 year aggregation.

Population

Routinely generated population estimates are provided from such entities as the National Center for Health Statistics (NCHS), but only on a county-wide level. Some jurisdictions may have access to local population small area estimates, which could be used for this project. Workgroup participants who did not have access to sub-county population estimates used 2010 Census data as a mid-point for LE. This is still a recommendation for small area estimates, if there are no other local or updated sources for the data.

The American Community Survey (ACS), the Census Bureau's annual household survey, does produce demographic characteristic distribution at a census tract level but is not recommended for use as a population denominator. The primary purpose of the ACS is to measure changes in a community's socioeconomic characteristics based on a small sample of households surveyed every month. The Census recommends ACS for gauging trends over time and for comparing characteristics across areas, but specify that it lacks the precision for population estimates. Many of the tracts have very high coefficients of variation, which indicate the lack of precision. In addition, ACS provides summary data for aggregated age groups, including 0-4; they do not include results for the <1 population. One Phase I Workgroup member compared their LE results using the 2010 Census population vs the 5-year ACS data, and showed that the ACS data markedly changed results.[10]

5.3 Methods Selected for SCALE

The Phase I SCALE Workgroup chose the adjusted Chiang II method because 1) it adjusts for the variance in the final age band and 2) it addresses age bands with zero deaths. Research shows simulation results suggested that use of the adjusted Chiang II method provides the closest approximations to reference LE and standard errors by not imputing any values into age groups with zero deaths [1], with the exception of the oldest age group.

An environmental scan of existing LE tools by the Workgroup led to discovery of an existing Life Expectancy Tool, created by the South East England Public Health Observatory (SEPHO) group (http://www.sepho.org.uk/viewResource.aspx?id=6626). This LE tool, which also uses an adjusted Chiang II method, is well documented. The Workgroup extensively tested and validated the results against SAS® (SAS Institute Inc. Cary, NC, USA.) programming used by LACDPH Stata (StataCorp. 2013. Stata Statistical Software: Release 13. College Station, TX, USA: StataCorp LP.) programs used by PHSKC, and an in-house Excel tool for calculation of LE created by NYS. Given the extensive

[10] Personal communication, SCALE Workgroup. Data to be released after paper publication.

documentation of the SEPHO tool, similarities of, ease of use, and accessibility, the Workgroup unanimously recommended it as a tool for SCALE. SAS and Stata tools are available upon request.

6.0. Acquiring and formatting data for SCALE

6.1. Acquiring Data

For jurisdictions that have not been involved in previous efforts to perform small-area analyses, it is not uncommon to spend several months (between six to nine months) acquiring data for the purpose of calculating sub-county LE estimates. One jurisdiction in Phase I needed to have an Institutional Review Board review their request for data. Connecting with the data providers and the stakeholders who could use the small area LE is beneficial in helping to gain access to the data as well as potential technical assistance with various aspects of the process, including cleaning and formatting data, geocoding, selecting an appropriate small-area, and statistical analysis.

6.1.1. Data sources, necessary variables, and formatting

Two types of data are needed: (1) death certificate data and (2) population estimates. The death certificate data and the population data will need to be generated at the same geographic level (census tract, ZIP code, city). Most **death certificate** data will not come assigned to a census tract or neighborhood level, but more state and local health departments are starting to geocode. See Section XX for considerations about which geographic area might be optimal for use. One of the most essential decisions involves the level of geography to present the data. Phase I participants chose census tract, aggregations of census tracts, ZIP code, and a Minor Civil Division. There are different factors driving the decision – for example, Florida chose to use ZIP code as they have a number of other health indicators also mapped at the ZIP code level. Maine tried a variety of geographies until they hit one that worked for their population and data reliability requirements. New York State used aggregations of census tracts (and suppression of tracts that were primarily group quarters). Several of the jurisdictions chose to use census tract for ease of incorporating social determinants of health from the American Community Survey.

Phase I and Phase II Workgroup participants typically used between 5 and 10 years of mortality data to be able to generate small area estimates. For those who used 5 years of data, they requested 2008-2012 data.

Population data may come from internally created numbers, or from the Census. Many Phase I participants began with the 5 year period spanning the 2010 Census, multiplying the Censal counts times 5 to generate a population estimate. A few had their own locally generated population estimates.

Both data sources should be broken down into 19 age groups (<1, 1–4, 5–9, 10–14, 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79, 80–84, 85+) for the purpose of this analysis. This will be fed into an Excel spreadsheet tool, so the age group aggregation should follow this set-up.

6.1.2. Formal data agreements and approvals to consider

Mortality data can be obtained from a statewide department of health[11] or a vital registration office.[12] State limitations, statutes, and practices differ with respect to the requirements for acquiring mortality data however in some cases this includes a requirement for a data sharing agreement (DSA) or a memorandum of understanding (MOU). In rare instances, some jurisdictions may also need to gain the approval of an Institutional Review Board (IRB). Localities new to the project may want to first do a quick internet or intranet search to see if previous work has already been created such as an MOU or DSA used by another county that can be repurposed. See Appendix B for examples of MOUs or DSAs used by SCALE participants.

It is important to forge relationships to help with data sharing. Analysts will need to make a number of decisions about how to calculate and present the data, and having stakeholder involvement in this process. See the Figure 12 for a flowchart summary walkthrough of suggested steps. At a minimum, **date of birth, gender, date of death, cause of death (primary and underlying), street address, ZIP code, city, and death certificate number** should be requested. Data users might want to consider asking for geocode (tract or latitude/longitude), a measure of geocode quality, and other demographic or socioeconomic pieces for additional analysis, such as race/ethnicity, occupation, etc. Date of birth rather than an age or age range is useful, especially for the <1 age group. Phase I and Phase II participants reported having had <1 "age" be recorded as a month, which the tools converted to a year, making calculations incorrect.

Some of these variables are not used to calculate LE but may be useful in pursuing analysis of LE disparities. Including them in the initial data request prevents needing to go back for another request. It's unlikely that jurisdictions will be able to use these direct estimates to examine the granular cause of death in small areas, but some Phase II participants were able to show impacts at a county or state level by collapsing categories of death (e.g. all cancers, all heart disease) and to look at population and geographical patterns.

Another question to address when requesting data from a vital stats provider is to know whether they have a reciprocal agreement to get data for residents of their state that die in another state. (See Section 8.2 – border areas for more discussion). If they do not, jurisdictions will need to decide if they want to consider approaching the other states, as this process may be time consuming and would be best initiated at the outset of the project.

One caveat: The standard death certificate form changed in 2004, and jurisdictions adopted it at different times. To date (2017), not all jurisdictions have adopted the new certificate, and some may have switched certificate forms during the 5-year period of interest. It's worth a discussion with the vital statistics registrar or others who work with the data to understand the impact of the certificate change on the data collected.

Once the data agreements are established (if necessary) and analysts have become more facile with the data, it allows for ease in multiple iterations of the process.

[11] https://www.cdc.gov/stltpublichealth/sitesgovernance/index.html
[12] https://www.cdc.gov/nchs/nvss/deaths.htm

Questions to consider when acquiring data for sub-county life expectancy

- √ What elements can be included without a DSA or MOU?
- √ Has another jurisdiction or organizational unit created a successful DSA or MOU that can be adapted?
- √ Are the data geocoded?
- √ Have other programs or units worked with geocoded mortality data? Might you be able to leverage this work?
- √ Will you need to work with other jurisdictions to get information about geographies along the border of your geography, such as counties bordering another state or a ZIP code that crosses county or state boundaries?
- √ Are there existing requirements around suppression or censorship of unreliable numbers?
- √ How many years of data might be needed for sufficient sample size?
- √ Could this data benefit other priorities in your LHJ?
- √ Will you want other demographic information for analysis (e.g. race/ethnicity, education, occupation)?

6.2. Geocoding Mortality Data

Geocoding is the process by which descriptors (e.g., address, city, ZIP code, province) are assigned a place on a map, also known as a geospatial location, creating a georeferenced dataset. Although this section is not designed to be a comprehensive tutorial, it will provide some links for additional information and best practices.[13] Recent assessments of state health department epidemiology capacity indicate that more than 50% do not routinely geocode their data.[30] As a result, it will be important many jurisdictions that wish to calculate sub-county LE will need to geocode their mortality data as part of this process. In this section we offer some tips to make this process smoother than it may be otherwise.

6.2.1. Geocoding defined

The Workgroup suggests the initial task of geocoding is to assign address information from a death certificate to a 2010 census tract. Because this project is examining geographic disparities in life expectancy, geocoding data is essential to examination of sub-county LE estimates.

Some jurisdictions may already have access to geocoded data when the data are requested; others may need to perform this step. Even when geocoded data are available, it is important to understand how the geocoding process was accomplished, and the attendant levels of accuracy resulting from the method used. For example, some automatic geocoders might assign a centroid (a center point) of a city or a ZIP code to an address that cannot be matched to a street location. This can artificially inflate the number of deaths that are occurring in that tract, as the individuals are being assigned to the numerator but are not in the denominator. If data come already geocoded, one recommendation

[13] http://naaccr.org/LinkClick.aspx?fileticket=ZKekacM8k_IQ0%3d&tabid=239&mid=699

would be to request a match score and match type, which will provide the data user with appropriate information in deciding how to use the data. The user can manually match the centroid or missing values, they can assign based on an overlay, or they can randomly distribute non-matching cases.

If data do span a Censal year, data users should also verify that any geocoded data are assigned to the same census geography. One Phase I participant reported needing to re-geocode the 2008 and 2009 data to 2010 Census boundaries.

6.2.2. Software for batch geocoding

Should a participant need to geocode the death data, there are several different software options available for batch geocoding. Batch geocoding occurs when the datasets are processed through a software package and assigned to a geolocation automatically. Any geographic information system (GIS) will have batch geocoding options. Following are several software options that may be valuable to explore in more detail:

- **ArcGIS (from ESRI).** . States that participate in the CDC Building GIS Capacity for Chronic Disease Surveillance (http://www.cdc.gov/dhdsp/programs/gis_training/index.htm) should have access to this software. A geocoding tutorial for this software is accessible on the Web: http://help.arcgis.com/en/arcgisdesktop/10.0/pdf/geocoding-tutorial.pdf.
- **Google Earth.** Google Earth has geocoding capability. To use the full functionality of this tool, it is helpful for the user to know HTML and JavaScript. Users must sign up for Google Earth and obtain an API key, which is what allows a user to connect and geocode using Google Earth. A tutorial is available. There is a paid, unlimited version and a free version, which is limited to 2,500 geocodes a day. http://www.drew.edu/ess/wp-content/uploads/sites/82/Tutorial-7-Geocoding-with-Batch-Geocode-and-Google-Earth.pdf
- **Statistical analysis packages.** Some analytic software packages, such as SAS and R, also have geocoding tools. For example, SAS/GRAPH uses a proc geocode (http://support.sas.com/documentation/cdl/en/graphref/63022/HTML/default/viewer.htm#a003121448.htm) and R has a package on CRAN http://www.inside-r.org/packages/cran/ggmap/docs/geocode. The default R package uses the Google Earth API, but other packages also can be used.

6.2.3. After batch geocoding

No software can match all addresses to a location, as there are errors in the address file, the street file for matching, or addition of news roads that aren't yet incorporated into the GIS software. Records that don't geocode are called exceptions. Exceptions can result from an incorrect street number, misspelling of a street, incorrect post office box, name of building instead of street, incorrect directional mismatch between ZIP code and street, or an incorrect underlying street layer as some examples. Rural areas may have route boxes that also are difficult to assign. These exceptions should be manually reviewed and matched when possible, as most exceptions do not happen randomly. Sometimes local GIS addresses will be better than a national one; some may just require spelling corrections. If a ZIP code is part of the record, and that ZIP falls entirely within a tract, that tract can be assigned to the record. A final geocoding match of $\geq$90% is ideal, and the higher it can be, the better. Most Phase I participants achieved a greater than 95% match score.

Addresses that remain unmatched after review can be handled one of two ways. They can be treated as missing, if there is non-differential classification (i.e., if the unmatched cases are similar to the matched cases in terms of demographics), they can be randomly assigned by having them distributed across all tracts, or geocodes can be imputed, based on demographic factor similarity between the case and the population.

6.3. Preparing the data

The population and death data need to be arranged in the 19 age groups listed above. This fits the abridged life table format. See Figures A and B as examples. Remember, if using the population estimates from the Census, multiple by 5 (or the number of years used to aggregate).

Figure A: Example of numerator/death data, 5 year aggregate[14]

GeoID	<1	1-4	5-9	10-14	15-19	20-24	25-29	...	70-74	75-79	80-84	85+
G01	32	39	3	3	11	13	22	...	141	163	196	488
G02	41	13	6	6	18	31	43	...	190	290	228	514
G03	27	32	2	2	2	2	16	...	145	161	194	518
G04	39	6	3	3	7	16	33	...	196	261	265	588

Figure B. Example of denominator/population data; 2010 population multiplied by 5

GeoID	<1	1-4	5-9	10-14	15-19	20-24	25-29	...	70-74	75-79	80-84	85+
G01	9005	10805	35125	42150	33890	40670	54140	...	25315	14715	17660	4810
G02	7865	25190	35330	29655	40070	36795	47215	...	19440	25335	11475	5300
G03	5840	7010	25820	30985	25845	31010	45860	...	32050	18260	21910	4915
G04	8780	27305	36505	27210	36635	39430	54615	...	18330	19005	9990	5665

There should be two spreadsheets: one each for the number of deaths and the number of population. In each spreadsheet, each row represents a small area, and each column represents an age group.

[14] Calculating LE for Small Areas, T. Talbot, CSTE 2016

Each row should also have a geographic identifier (city, ZIP, census tract). Those using the SEPHO tool described in Section 7 will see this referred to as an **area code**. This geographic identifier is required to calculate LE for individual areas.

Questions to consider with respect to preparing the data

- √ Are there geographic areas with populations less than 5000? Users may want to flag these areas for evaluation of the result.
- √ Are there cells with no deaths in the <1 age category or 85+?
- √ What does the distribution of deaths look like across tracts? Might tracts need to be grouped together for reliable results?
- √ Will I need to request data from another jurisdiction that abuts my boundaries?
- √ What is the quality of the geocodes?

7.0. Selecting a Software: Available Options

Workgroup participants scanned the available toolsets, and this Guide presents the 3 most common options. All Phase I and Phase II members used the SEPHO tool (See Section 5 for more information), from the South East England Public Health Observatory group, who have been developing small area LE for many years. This tool was verified and tested for accuracy. Given the simple nature of needing no special statistical software, the ease of cutting and pasting, this tool has a low bar to entry for ease of use. It also has many sub-tabs that allow the user to see exactly what is going on in each calculation and can be useful in helping to discover why some results may be different than expected.
http://webarchive.nationalarchives.gov.uk/20160701122411/http://www.sepho.org.uk/viewresource.aspx?id=8943

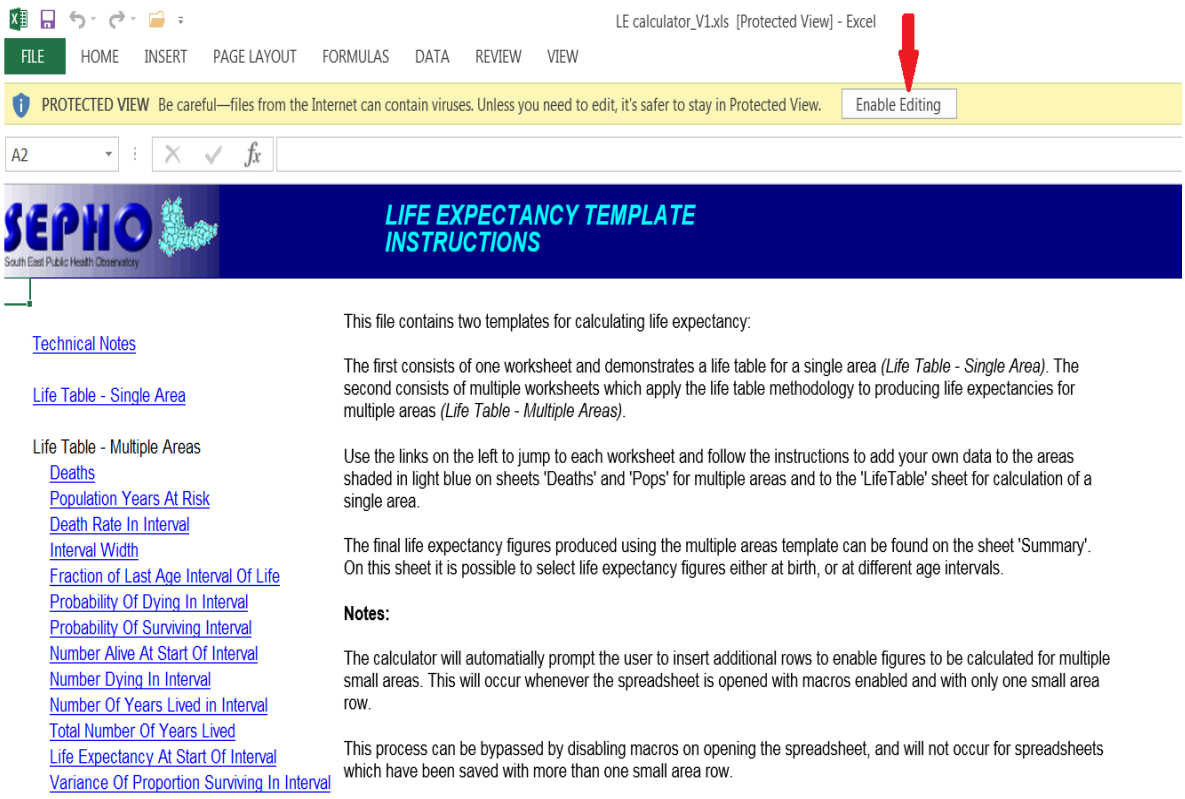
7.1 Using the SEPHO Excel tool

The numerator and population data will be copied and pasted in the SEPHO Excel tool, which contains macros that will calculate the result. This is a simple, easy to use tool that was verified for accuracy against statistical programs. (See section 7.2 for more details). This will involve a download from a website, and works best with Excel version 10 or higher. Phase I and Phase II SCALE participants used the SEPHO tool for simplicity. Regardless of the tool used, the steps in preparing data are the same for the calculations.

SEPHO tool Download

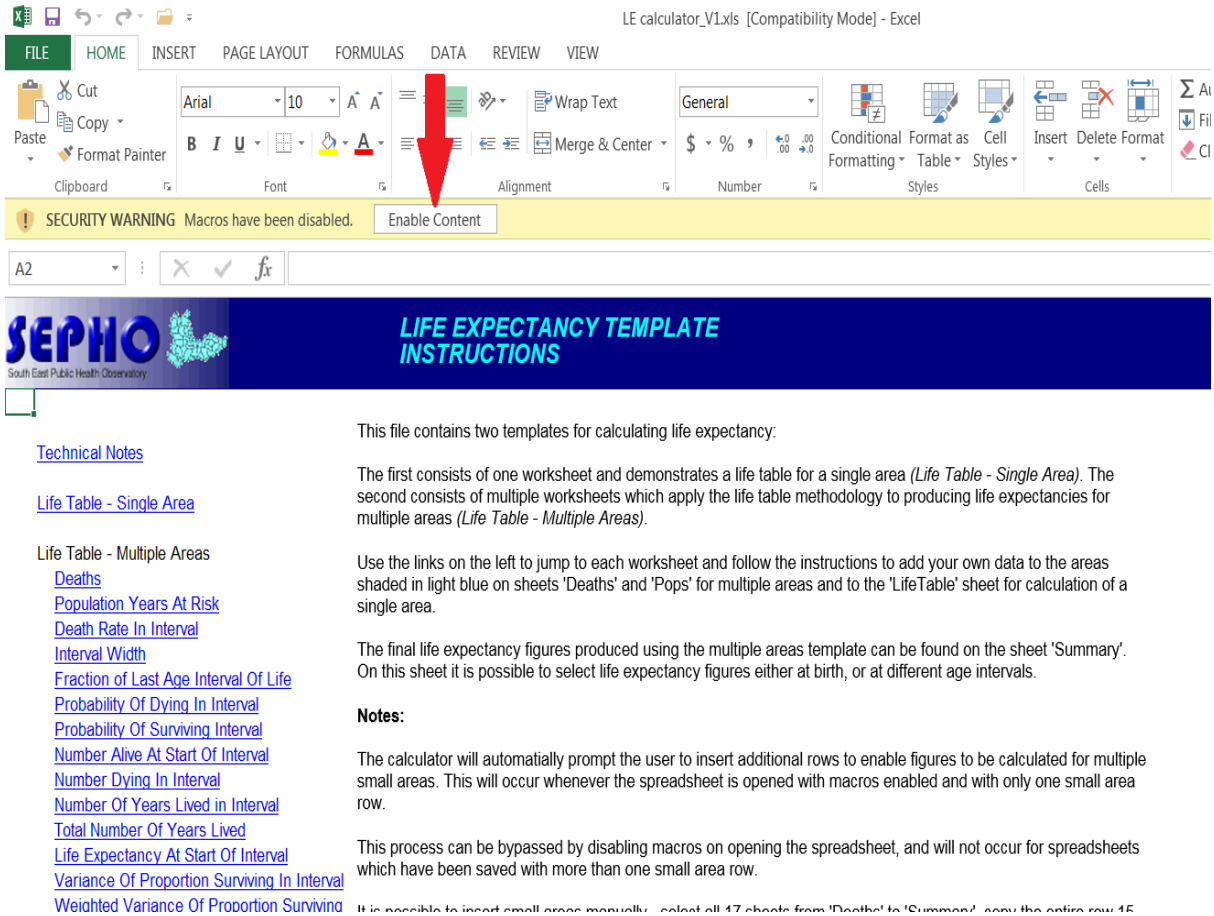
1. Download the SEPHO tool from http://webarchive.nationalarchives.gov.uk/20160701122411/http://www.sepho.org.uk/viewresource.aspx?id=8943
2. Open the file (Life Expectancy calculator_V1.xls)
3. If this is the first time you are opening the file, depending on the security setting of your Microsoft Office, you might need to click “Enable Editing” (see Figure 8 for an example)

Figure 8. SEPHO tool, highlighting the “Enable Editing” button.



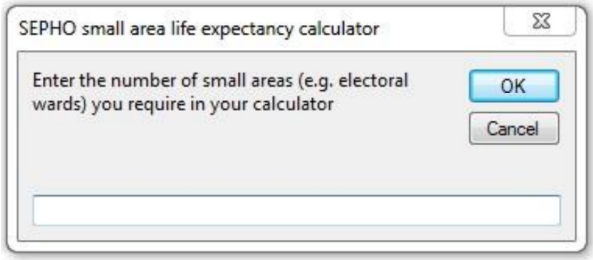
4. Again, depending on the security settings of your Microsoft Excel, you might need to enable the Macro by clicking “Enable Content” (where the red arrow points).

Figure 9. SEPHO tool, showing where to click to enable the macro.



5. Suggest saving the enabled version to have an unedited copy.
6. Click on “Deaths” under Life Table – Single Area.
7. A dialog box titled “SEPHO small area life expectancy calculator” will pop out requesting that the user “[E]nter the number of small areas (e.g., electoral wards) you require in your calculator.” Put in the number of small areas in the box below – this should match the number of geographic units (tract, ZIP MCD, etc) being used in the calculation. Make sure your data are sorted by geography.

   Figure 10. SEPHO calculator pop up



8. Copy the numbers of death by age group to the “Deaths” spreadsheet (if the spreadsheet is not shown, click the right arrow on the bottom left corner to show the “Deaths” spreadsheet).Copy the number of population by age group to the “Pops.” Data should be sorted by geography so that the order matches in the Deaths and Pops page.

Figure 11. Spreadsheet example.

Area Code	Area Name	Age Group																			Total
		Under 1	1-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+	All Ages
Code	Standard/Nat	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Code	Small area	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Note that the column "Area code" is required for calculating life expectancy and standard errors at the individual area level. This is just the geographic identifier from the input data set. If the area code is not provided, the life expectancy and standard error will be calculated for the aggregation but not for the individual areas.

9. Results will be displayed on the spreadsheet "Summary." Intermediate results are displayed as well with Life Expectancy at Start of Age interval displayed on spreadsheet "e." Make sure the table is set to "Birth" to compute LE at birth. Users can also change this to look at the impact of mortality rates on other age groups as well, but that is beyond the scope of the SCALE Guide.
10. To calculate life table for a single area, copy and paste your data on to the spreadsheet "LifeTable," results will be displayed on columns "P, U,V,W" for the point estimate of the life expectancy at the start of the age interval, sample standard error, lower bound of the 95% confidence interval, and upper bound of the 95% confidence interval.
11. Save the Excel Workbook under a different name.
12. Sophisticated users can modify the life table tab if they prefer to use a different life table.

More details are available at
http://webarchive.nationalarchives.gov.uk/20160701122411/http://www.sepho.org.uk/viewresource.aspx?id=8943

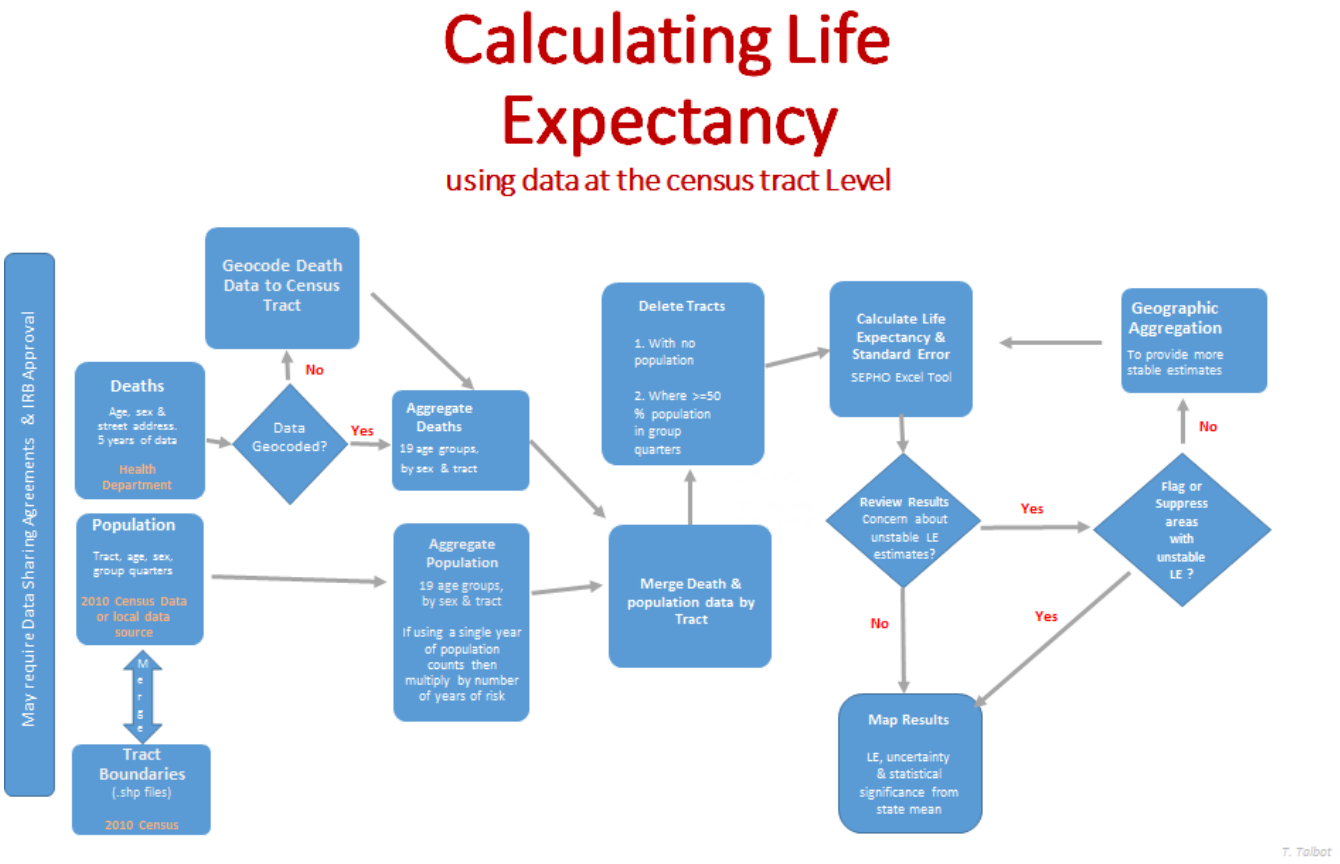
7.2 SAS or Stata option

For a jurisdiction wishing to use a SAS- or Stata-based tool for calculating life expectancy, CDC has a SharePoint site hosting language used for testing purposes and for which limited technical assistance is available. Numerator and denominator data need to be created as described above for the SEPHO tool.

7.3 Flowchart

One of the Phase I Workgroup members created a flow chart to provide a visual giving the workflow. Most Workgroup members needed to run through several iterations to reach a final product that was ready for analysis, display, and distribution.

Figure 12. A flowchart of activities for calculating LE.



8.0 Interpreting the Findings

8.1 Phase I results

Of the Phase I participants, four used death data from 2008-2012; two used more recent data (2009-2013), and one jurisdiction with more sparse populations used 10 years of data (2001-2010) to generate sub-county LE. Four jurisdictions used Census population estimates while three had local population estimates. All participants found the SEPHO tool easy to use, and customizable for the specific type of geographies pertinent to their locality.

Each jurisdiction examined the output of the numerator, denominator, and the standard errors, and made some decisions about what to present. One Phase I participant (Florida) initially used ZIP codes as it aligned with other sub-county health and behavioral indicators that were already produced. Maine needed to group 10 years of data and provide them at the Minor Civil Division (MCD), which is a Census designation for a civil township, precinct, or magisterial district. The other jurisdictions produced census tract estimates.

Many jurisdictions found that mapping the data after the first pass through the SEPHO tool led to ask more questions (see section 8.2). The Workgroup discussed sensitivity analyses, what to consider unreliable estimates, how to handle unreliable estimates, and special considerations around specific geographic types.

8.2 Special considerations

Areas with unexpectedly high or low LE

Border areas

One question to ask is if the vital statistics department has a reciprocal agreement with other states. This means that a resident of State A dies in another state, State B, that State B would send a death certificate to State A. Not having this agreement means that LE might be artificially inflated in areas that are close to borders, especially ones that have a hospital or nursing facility, as individuals might be seen in the hospital or nursing facility in State B, die there, and State A would not know. Border issues typically arise between county, country, and state lines.

Small populations:

In some cases, tracts may not meet the recommended minimum size of 5,000 person years at risk due to a small number of people residing in that tract. Others may have only a few deaths. See section unique tracts below for tracts that have unique characteristics causing the small population. Jurisdictions participating in Phase I and Phase II decided to handle small populations in one of several ways. Some suppressed the data. Some calculated the LE, and if it had a reasonable standard error, the data were presented. Others grouped additional years, other geographies, and some aggregated geographies together. NYS aggregated tracts to get to 60 deaths, using their R-based Geographic Aggregation Tool. [15]

Unique tracts:

Some tracts may contain a hospital or nursing facility. It's possible that the facility could be put on the death certificate as the residence, causing a spike in the number of deaths in that tract.

Many tracts have zero or very small population; for example, many urban areas have assigned a tract specific to an airport or to a prison. One jurisdiction identified these using data from the Census or ACS, excluding tracts that have more than 50% of the population living in group quarters. Deaths can still occur in those geographies, but the individuals would have an official residence somewhere else.

Standard errors and confidence intervals

Standard error (and hence confidence limits) increase as population decreases. Literature suggests that a population of 5,000 life years at risk produces an 'acceptable' standard error of +/– 2 years (or a 95% confidence limit of +/– 4 years). Phase I and Phase II participants found that even with 5,000 person years at risk, there were still small areas with a standard error outside the range of 2. How to handle these areas is still an item up for discussion.

[15] http://www.albany.edu/faculty/ttalbot/GAT/

Impact of Migration on LE

Populations that may be highly mobile or areas that have large numbers of influxing younger, healthy people, may see changes in life expectancy.

8.3 Limitations of the tool

9.0 Using the LE estimates

Policy makers and community members often ask what is driving the geographic disparities that are seen. The National Environmental Public Health Tracking Network (NEPHT) has produced list of useful related social-economic indicators which are important when looking at health outcomes such as LE. Virginia also computed Healthy Life Expectancy. As Phase I and Phase II participants release data, this Guide will be updated.

10.0 Mapping and Display of LE results

Phase I and Phase II participants who had successfully generated LE were surveyed in August 2016, to get a sense of how they were choosing to display their results. Most participants were using static maps generated in mapping software, although a few had some interactive maps on the internet. Some participants were mapping other social determinants of health alongside the LE estimates, including education, poverty, lack of health insurance, and risk factor or behavioral data.

Even in creating static maps, no gold standard has yet emerged on the best way to present the data. All participants created a classified thematic or choropleth map where shades represent a range of values. Each tract falls into a specific “bin” or range. There was a wide variety of methods in how the map cutpoints were created.

Equal interval:

Each class consists of an equal data interval along the dispersion (from the high to the low point) of the data. Intervals are determined by dividing the range of all your data by the number of classes desired. Equal intervals are recommended if the data is distributed in a rectangular shape or if classification steps are nearly equal in size. The major disadvantage of this method is that class limits fail to reveal the distribution of the data along the number line. There may be classes that remain blank, which of course is not particularly meaningful on a map.

Natural Breaks:

Groups (or clusters) the data around different classes, with the goal of minimizing the deviation in each class while maximizing the difference between the other groups. In other words, it minimizes value differences between the data within the same class and emphasize the differences between the classes.

A disadvantage of this method is that class limits may vary from one map-maker to another due to the author's subjective class definition (Good graphic way of determining natural group of similar values by searching for significant depressions in frequency distribution. Minor troughs can be misleading and may yield poorly defined class boundaries.

Quantiles:

This is a rank order approach, using an equal number of observations into each class, generated by taking the total observations and dividing by the desired number of classes. Each class is approximately equally represented. However, it may introduce an apparent pattern of dissimilarity if there is none, overweight the data, or there may be gaps between the observations.

Standard Deviation:

This method measure how the value is distributed along a dispersion graph, and it show the standard deviation from the statistical mean of our dataset. The resulting classes reveal the frequency of elements in each class. Standard deviation may be useful to show statistical significance or to direct attention to the high and low valued. It's typically best for a standard normal distribution. One downside is it doesn't list the exact LE rate, and may be difficult for a layperson to understand.

Table 2.

Methods employed to generate legends (n=22)	
Natural breaks Jenks	6
Quantile	4
Defined interval	1
Standard deviation	1
Geometric interval	1
Other*	3

*included manual classification into bins and difference from statewide average.

Another display/mapping issue includes determining how to represent unstable rates. Some chose to grey out or crosshatch unreliable data. Some suppressed; others aggregated sub-county units until they did not need to suppress the data.

Table 3.

Methods employed to indicate areas with high SE (n=12)*	
Greying out	5
Crosshatch	3
Suppressed	3
Aggregated	2

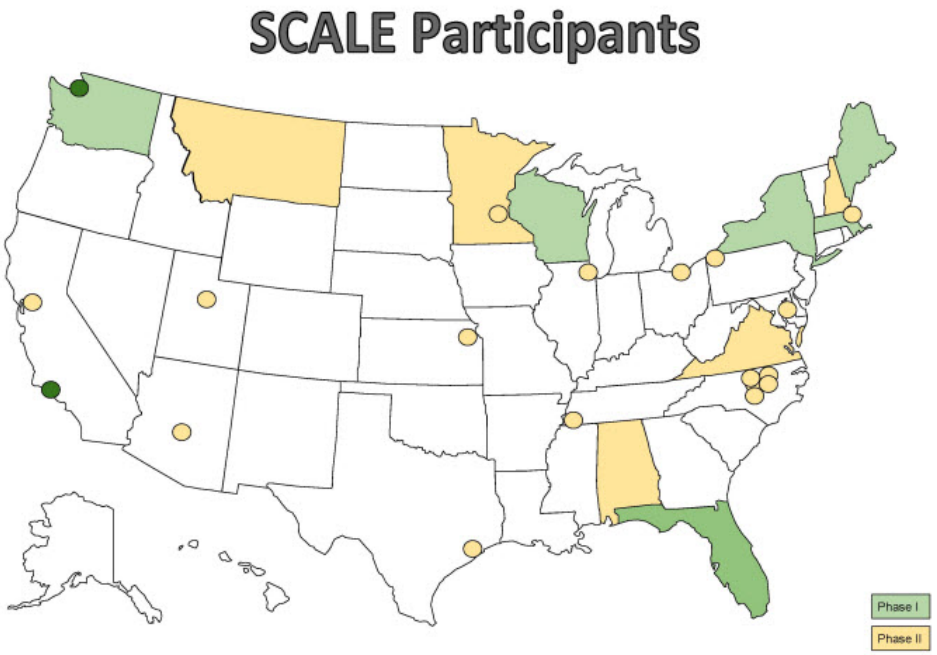
*Participants indicated multiple methods were employed

11.0 Summary

The SCALE project successfully piloted a simple way for local jurisdictions to calculate LE at sub-county areas. There are many lessons to be learned from the process, including how to obtain the correct data, using an appropriate tool, and the power of collaboration. Many of the Phase I and Phase II participants have been able to use their LE results locally and several have papers about the process in

the works. As those are released, the Guide and SCALE page will be updated. Feedback about the guide or questions about the project may be directed to CSTE.
http://www.cste.org/?page=SCALE&hhSearchTerms=%22scale%22

12.0 Acknowledgements



SCALE Phase I Participants

Florida
Los Angeles County, CA
Maine
Massachusetts
Washington
New York
Seattle & King County. WA
Wisconsin

SCALE Phase II Participants

Alabama
Maricopa Co., AZ
Alameda Co., CA
District of Columbia
Cook Co., IL
Alamance Co., NC
Caswell Co., NC
Chatham Co., NC
Durham Co., NC
Orange Co., NC
Johnson Co., KS
Erie Co., PA
Shelby Co., TN
Houston, TX
Montana
New Hampshire
Cleveland, OH
Metro Area Planning Council, MA
Washington Co., MN
Minnesota
Salt Lake Co., UT
Virginia

References

1. Toson, B. and A. Baker, *Life expectancy at birth: methodological options for small populations.* National statistics methodological series, 2003. **33**.
2. Los Angeles Department of Public Health, O.o.H.A.a.E., *Life Expectancy in LA County: How long do we live and why? A cities and community report.* . 2010.
3. Silcocks, P.B.S., D.A. Jenner, and R. Reza, *Life expectancy as a summary of mortality in a population: statistical considerations and suitability for use by health authorities.* Journal of Epidemiology and Community Health, 2001. **55**(1): p. 38-43.
4. Jonker, M.F., et al., *The impact of nursing homes on small-area life expectancies.* Health & Place, 2013. **19**(0): p. 25-32.
5. Harper, S., R. MacLehose, and J. Kaufman, *Trends in the black-white life expectancy gap among US states, 1990-2009.* Health affairs, 2014. **33**(8): p. 1375-82.
6. Adam Smith, J., S. Harper, and N. Auger, *Causes of widening life expectancy inequalities in Québec, Canada, 1989-2004.* Canadian Journal of Public Health, 2011. **102**(5): p. 375-81.
7. Stephens, A.S., et al., *Life expectancy estimation in small administrative areas with non-uniform population sizes: application to Australian New South Wales local government areas.* BMJ open, 2013. **3**(12): p. e003710-e003710.
8. Wang, H., et al., *Age-specific and sex-specific mortality in 187 countries, 1970-2010: a systematic analysis for the Global Burden of Disease Study 2010.* Lancet (London, England), 2012. **380**(9859): p. 2071-94.
9. Avendano, M. and I. Kawachi, *Why do Americans have shorter life expectancy and worse health than do people in other high-income countries?* Annual review of public health, 2014. **35**: p. 307-25.
10. Syme, S.L., *Reducing racial and social-class inequalities in health: the need for a new approach.* Health affairs, 2008. **27**(2): p. 456-9.
11. Banks, J., et al., *Disease and disadvantage in the United States and in England.* JAMA: the Journal of the American Medical Association, 2006. **295**(17): p. 2037-45.
12. Bezruchka, S., *The Hurrider I Go the Behinder I Get: The Deteriorating International Ranking of U.S. Health Status.* Annual review of public health, 2012. **33**(1): p. 157-173.
13. Squires, D., *The global slowdown in health care spending growth.* JAMA: the Journal of the American Medical Association, 2014. **312**(5): p. 485-6.
14. Woolf, S.H. and P. Braveman, *Where health disparities begin: the role of social and economic determinants--and why current policies may make matters worse.* Health Affairs, 2011. **30**(10): p. 1852-9.
15. Ezzati, M., et al., *The reversal of fortunes: trends in county mortality and cross-county mortality disparities in the United States.* PLoS Medicine, 2008. **5**(4): p. e66.
16. Murray, C.J.L., et al., *Eight Americas: investigating mortality disparities across races, counties, and race-counties in the United States.* PLoS Medicine, 2006. **3**(9): p. e260.
17. Cullen, M., C. Cummins, and V. Fuchs, *Geographic and racial variation in premature mortality in the U.S.: analyzing the disparities.* PLoS ONE, 2012. **7**(4): p. e32930.
18. Wang, H., et al., *Left behind: widening disparities for males and females in US county life expectancy, 1985-2010.* Population Health Metrics, 2013. **11**(1): p. 8.
19. *Measuring the Risks and Causes of Premature Death: Summary of Workshops*. 2015, Washington DC: 2015 by the National Academy of Sciences.
20. Hunt, B., G. Tran, and S. Whitman, *Life Expectancy Varies in Local Communities in Chicago: Racial and Spatial Disparities and Correlates.* Journal of Racial and Ethnic Health Disparities, 2015: p. 1-9.

21. Chen, J., et al., *Mapping and measuring social disparities in premature mortality: the impact of census tract poverty within and across Boston neighborhoods, 1999-2001.* Journal of urban health, 2006. **83**(6): p. 1063-84.
22. King County Equity and Social Justice Annual Report, November 2014. http://www.kingcounty.gov/elected/executive/equity-social-justice.aspx, last checked 10/2015.
23. Song L, Mercer L, Wakefield J, Laurent A, Solet D. Using Small-Area Estimation to Calculate the Prevalence of Smoking by Subcounty Geographic Areas in King County, Washington, Behavioral Risk Factor Surveillance System, 2009–2013. Prev Chronic Dis 2016;13:150536.
24. Chiang, C.L., *A stochastic study of the life table and its applications. II. Sample variance of the observed expectation of life and other biometric functions.* Human biology, 1960. **32**: p. 221-238.
25. Bravo, JM and Malta, J. Estimating life expectancy in small population areas. United Nations Statistical Commission and Economic Commission for Europe. 23 April 2010.
26. Arias, E., *United States life tables, 2008.* National vital statistics reports, 2012. **61**(3): p. 1-64.
27. Chang, C.L., *Introduction to stochastic processes in biostatistics.* 1968.
28. Eayres, D. and E.S. Williams, *Evaluation of methodologies for small area life expectancy estimation.* Journal of epidemiology and community health, 2004. **58**(3): p. 243-9.
29. Scherbov, S. and D. Ediev, *Significance of life table estimates for small populations: Simulation-based study of estimation errors.* Demographic Research, 2011. **24**: p. 527-550.
30. Hadler, J.I. Lampkins, R., Lemmings, J., et al. Assessment of Epidemiology Capacity in State Health Departments—United States, 2013. MMWR Morb Mortal Wkly Rep 2015;64: 394-398.

Appendix A Summary of Peer-Reviewed Literature (to be added)

Appendix B Examples of MOU/DSA (TBA)