Technical Notes: Methodology for Developing the Local Public Health Resources Index (LPHRI)

October 2024

Suggested Citation: Heluna Health. (2024). Technical Notes: Methodology for Developing the Local Public Health Resources Index (LPHRI). Heluna Health.

Introduction and Framework

The Local Public Health Resources Index (LPHRI) focuses on local public health resources needed to implement outbreak preparedness and response activities. The COVID-19 pandemic spurred interest in exposed the need to rebuild and transform public health infrastructure, including bolstering the public health workforce and improving public health data systems [1-3]. In 2022, the U.S. Centers for Disease Control and Prevention (CDC) released a Public Health Infrastructure Grant (PHIG), a multi-billion dollar investment that focuses on three key strategic areas: workforce, foundational capabilities, and data modernization [4]. Additionally, funding for public health continues to be an important impactful factor, with prior research linking local public health expenditures to declines in preventable causes of death [5].

To our knowledge, aside from surveys of samples of local public health departments, no comprehensive database of local public health resources in the U.S. exists. Therefore, we sought to bring together local-level data on these topics. Recognizing the importance of these categories of preparedness, and based on feedback from technical advisors, we created the following framework for the index:

- Workforce Capacity: what workforce is currently available in the public and private sector related to public health preparedness?
- Data Modernization/Technology Innovations: to what extent have new methodologies or technologies been adopted to improve public health preparedness, including data infrastructure and biosurveillance capabilities?
- Public Health Expenditures: what are the total local public health expenditures in the geography?

Through a literature search and consultation with technical experts, we identified two major groupings of workers that are critical to public health preparedness. First, the local public health agency workforce is an essential component of preparedness and response, including:

- Epidemiologists and data analysts, who conduct surveillance, investigation, and reporting activities
- Environmental health workers, who enforce environmental regulations that reduce disease transmission (e.g. foodborne diseases, water-borne diseases, etc)
- Public health laboratory workers, who staff laboratories that test samples as part of biosurveillance activities
- Nurses, who provide nursing support to communicable disease programs
- Community health workers and health educators, who provide a critical linkage and disseminate health information to communities
- Emergency preparedness workers, who create public health emergency response plans and provide related training

Second, since clinical laboratory capacity is important in diagnosing diseases presented in clinical settings, we included two key categories of clinical laboratory workers – phlebotomists and clinical laboratory technicians and technologists [6].

Additionally, given the role of pharmacists in distributing medical countermeasures (MCMs), we included pharmacists in this workforce category, which primarily includes private sector workers.

Third, public health expenditures at the local health department level were included as the final domain in the LPHRI.

Finally, we identified wastewater surveillance as a relatively new technology that was adopted more widely during the COVID-19 pandemic [7]. Use of this technology and reporting of results to the CDC National Wastewater Surveillance System (NWSS) are reflected in these indicators.

Data sources and definitions

Based on the framework described above and availability of data, the following indicators were developed for the LPHRI. In addition to using the most recently available data for the LPHRI, we calculated a "prior version" using data that are approximately 1-2 years older than the latest version, so we can compare temporal trends in these indicators. Table 1 describes the data sources and definitions for each indicator.

Table 1. Indicators and Data Sources for the LPHRI

Domain	Indicator Name	Data Sources	Data Year(s) (current)	Data Year(s) (prior)	Description
Public Health Workforce	Epidemiology Staffing	California, Government Compensation data (2019-2022), Transparent California (2020, 2021) Nevada, Transparent Nevada (2019-2022) Arizona, Open Payrolls, govsalaries.com (2019-2022; note: not all years available for all LHDs) Utah, Transparent Utah: (2019-2022; note: not all years available for all LHDs) Denominator: American Community Survey (ACS) 5-year populations for corresponding years	2021, 2022	2019, 2020	Number of epidemiologists, data analysts, GIS analysts, research analysts in the local health department (LHD), per 100k population. Note: temporary contact tracers/disease investigators were not included in this category because these positions were not consistently reported across jurisdictions. Where maximum salaries were available in the dataset, the percent full time equivalent (FTE) was calculated before summing the total number of FTEs for this job category. The indicator is calculated by first calculating the number of staff per 100k population for each year, and then averaging over the 2 years.
Public Health Workforce	Nursing Staffing		2021, 2022	2019, 2020	Number of nurses employed in the LHD, per 100k population. This indicator includes all levels of nurses, such as registered nurses, nurse supervisors, licensed vocational nurses, and public health nurses. Where maximum salaries were available in the dataset, the percent full time equivalent (FTE) was calculated before summing the total number of FTEs for this job category. The indicator is calculated by first calculating the number of staff per 100k population for each year, and then averaging over the 2 years.
Public Health Workforce	Laboratory Staffing		2021, 2022	2019, 2020	Number of laboratory workers in the LHD, per 100k population. This indicator includes microbiologists, laboratory scientists, laboratory assistants, laboratory technicians, and related laboratory personnel. Where maximum salaries were available in the dataset, the percent full time equivalent (FTE) was calculated before summing the total number of FTEs for this job category. The indicator is calculated by first calculating the number of staff per 100k population for each year, and then averaging over the 2 years.
Public Health Workforce	Environmental Health Staffing		2021, 2022	2019, 2020	Number of environmental health workers in the LHD, per 100k population. This indicator includes environmental health specialists, environmental management personnel, environmental technicians, environmental inspectors, and related personnel. Where maximum salaries were available in the dataset, the percent full time equivalent (FTE) was calculated before summing the total number of FTEs for this job category. The indicator is calculated by first calculating the number of staff per 100k population for each year, and then averaging over the 2 years.
Public Health Workforce	Community Health Worker Staffing		2021, 2022	2019, 2020	Number of community health workers and health educators in the LHD, per 100k population. This indicator includes community health workers, community health technicians, community liaisons, community outreach workers, community service workers, health educators, peer counselors, and related personnel. Where maximum salaries were available in the dataset, the percent full time equivalent (FTE) was calculated before summing the total number of FTEs for this job category. The indicator is calculated by first calculating the number of staff per 100k population for each year, and then averaging over the 2 years.
Public Health Workforce	Public Health Emergency		2021, 2022	2019, 2020	Number of emergency preparedness staff in the LHD, per 100k population. This indicator includes emergency preparedness workers, emergency planners, preparedness specialists, preparedness coordinators, and related personnel.

	Preparedness Staffing				Where maximum salaries were available in the dataset, the percent full time equivalent (FTE) was calculated before summing the total number of FTEs for this job category. The indicator is calculated by first calculating the number of staff per 100k population for each year, and then averaging over the 2 years.
Clinical Lab and Pharmacy Workforce	Clinical Laboratory Technologists and Technicians	BLS Occupational Employment and Wage Statistics Denominator: ACS 5-year populations for corresponding year	2021, 2022	2019, 2020	Number of clinical laboratory technologists and technicians per 100k population. Clinical laboratory technologists and technicians help prepare and process biological samples for laboratory testing. Clinical laboratory workforce is an important component of disease surveillance. The indicator is calculated by first calculating the number of staff per 100k population for each year, and then averaging over the 2 years.
Clinical Lab and Pharmacy Workforce	Phlebotomists	BLS Occupational Employment and Wage Statistics Denominator: ACS 5-year populations for corresponding year	2021, 2022	2019, 2020	Number of phlebotomists per 100k population. Phlebotomists play an important role in drawing blood and collecting specimens for clinical laboratory testing. Clinical laboratory workforce is an important component of disease surveillance. The indicator is calculated by first calculating the number of staff per 100k population for each year, and then averaging over the 2 years.
Clinical Lab and Pharmacy Workforce	Pharmacists	National Plan and Provider Enumeration System (NPPES) Denominator: ACS 5-year populations for 2022 and 2021	NPPES Jan 2024, ACS 2022	NPPES Nov 2022, ACS 2021	Pharmacist workforce in the county, per capita. The pharmacy taxonomy codes included are: - Pharmacist: 183500000X o Ambulatory Care: 1835P2201X o Critical Care: 1835C0205X o Geriatric: 1835G0303X o Nuclear: 1835N0905X o Nutrition Support: 1835N1003X o Oncology: 1835X0200X o Pediatrics: 1835P0200X o Pharmacist Clinician/Clinical Pharmacy Specialist: 1835P0018X o Pharmacotherapy: 1835P1200X o Psychiatric: 1835P1300X - Pharmacy technician: 183700000X Pharmacists dispense medications (including vaccines) and, as such, are a crucial workforce component for distributing medical countermeasures in response to a disease outbreak.
Public Health Expenditures	Public Health Expenditures	Arizona: public health department expenditures abstracted from Schedule F in agency budget reports (FY19, 20, 21, 22) for all counties except Graham County, where records were downloaded from OpenBooks.az.gov, and Yuma County, where records were abstracted from the agency website. California: City and County agency expenditures reports downloaded from state controller's office (FY19, 20, 21, 22) Nevada: public health department expenditures abstracted from agency budget reports, FY19, 20, 21, 22	FY21, FY22	FY19, FY20	Total local public health agency expenditures per capita. This indicator was calculated by dividing the total public health expenditures, divided by the population denominator, and averaged over the two years. Local public health expenditures have been linked to improvements in preventable mortality. Arizona: Expenditures for Public Health were included. California: Expenditures categorized as "Public Health" were included Nevada: Expenditures for Public Health departments were included. Budget information for this time frame was not available for Central Nevada Health District, which was established in December 2022. For several counties that do not have their own local health department (and whose public health services are provided by the state of Nevada), we did not include budget information, as we would not be able to distinguish the portion of the state budget that serves these counties.

		Utah: Expenditure data were downloaded from Transparent Utah for the following health departments: Bear River, Central Utah, Davis County, Southeastern Utah, and TriCounty. Expenditures were abstracted from agency budget reports for Salt Lake, Summit, Tooele, Utah, and Wasatch counties, Weber-Morgan Health Department and Southwestern Utah Health Department (FY19, 20, 21, 22) ACS 5yr population estimates for the corresponding year were used as denominators			Utah: Expenditures for Health Districts or Public Health departments within county agencies were included.
Wastewater Surveillance	Wastewater Surveillance: population coverage	CDC National Wastewater Surveillance System, downloaded [Jan 10, 2024]	2023	2022	Population coverage for wastewater surveillance testing. This indicator was calculated as the proportion of the county population that resides within a sewer shed participating in wastewater surveillance testing. Wastewater surveillance can be an important piece of a county's monitoring of population-level disease trends. Counties served by sewer sheds with the infrastructure to participate in wastewater surveillance may be better prepared to identify and track disease outbreaks. For sewersheds that crossed county borders, we assigned these sewersheds to the county where the majority of the sewershed was located.
Wastewater Surveillance	Wastewater Surveillance: frequency of testing	CDC National Wastewater Surveillance System, downloaded [Jan 10, 2024]	2023	2022	Frequency of wastewater surveillance testing. This indicator was calculated by taking the total number of days of wastewater surveillance testing in a year divided by the total number of months that testing was done. More frequent testing allows for faster detection of a signal, which can provide an early indicator of a new disease outbreak.

Methods

1. Selection and aggregation of indicators

    1.1. Statistical methods to reduce dimensionality
    We used the Barnes-Hut implementation of the t-distributed Stochastic Neighbor Embedding (t-SNE) as an exploratory method for potential dimension reduction. While two distinct clusters were identified using t-SNE, we decided to keep all the initial indicators. For indicators using both the most recent data and the "prior version", there was a single cluster populated mostly with small counties and a second cluster with mostly large and medium sized counties. The analysis was conducted using Rtsne version 0.17.

    1.2. Aggregating Indicators

    Before combining individual indicators, the data were centered and scaled. Each composite domain score is the arithmetic mean of all individual indicators within that domain. Due to the use of multiple imputation, a composite score was computed on each of the ten complete datasets and the final composite score is the pooled result of all ten composite indices.

    1.3. Conversion of county-level data and local health jurisdiction (LHJ) data

    Because some data were available at the county level and other data were available at the local health jurisdiction (LHJ) level, it was necessary to convert the datasets to the appropriate geography. To aggregate data points from a smaller geography to a larger geography, the conversion was done through a population-weighted average of the smaller geographies to calculate the value for the larger geography. For example, for health districts that consist of multiple counties, the health district value for an indicator is calculated as the population-weighted average of the county-level indicator values. Conversely, when converting data from a larger geography to a smaller geography, such as when a county contains more than one LHJ, the LHJ's within that county will all be assigned the same indicator value as the county in which the LHJ is located.

    1.4. Weighting
    Due to the absence of convincing empirical evidence suggesting the use of unequal weights, we applied equal weights to all indicators thus allowing for equal contribution to the composite domain score. Expert opinions can also be used to guide weighting decisions and can be considered for future releases.

2. Diagnostics

    2.1. Missing Data

    There was some degree of missing data present in ten of the twelve indicators. The indicators calculated with the most recent available data had 149 missing observations across all twelve indicators. There were 39 incomplete records amounting to 32% of the data. The phlebotomy indicator had the most missing records with 23 total missing observations. The set of “prior version” indicators (i.e. using data that is not the most recent available) had a total of 282 missing observations across all twelve indicators, resulting in incomplete cases in a total of 62 records or 51% of the data. Of these missing records, 45 included the phlebotomy indicator. Multiple imputation with classification and regression trees (CART) via the MICE package (version 3.16.0) in R was used to obtain a complete dataset. Pooled results from the ten complete datasets produced with multiple imputation were used for all analyses. The data displayed on the webpage only uses imputed data for composite domain scores, while the display of individual indicator scores only use non-imputed data. A sensitivity analysis was conducted to assess the impact of multiple imputation and the method of imputation used. The imputation methods we considered were classification and regression trees, Bayesian linear regression, and predictive mean matching.

2.2. Indicator Distributions

The distributions of all individual indicators were evaluated. All wastewater indicators exhibited a strong degree of right skewness. To minimize undue influence on the composite index, the wastewater indices were log transformed. All other indicators were left untransformed.

2.3. Correlations

In order to obtain a complete correlation matrix to evaluate the correlations between all individual indicators, the imputed data were used. As expected, the two wastewater indicators are positively correlated with each other and the two clinical laboratory staffing indicators are positively correlated with each other. Similar degrees of correlation exist among the version using the most recent available data and the “prior version” of data.

3. Sensitivity Analysis

We compared the distribution of the final composite domain scores as produced by three different multiple imputation techniques: classification and regression trees (CART), Bayesian linear regression, and predictive mean matching, as well as no imputation. The final composite domain scores (latest available data and previous release) were robust to the method used to handle missing data.

4. Validation

We made a robust though not exhaustive effort to validate LPHRI domains and indicators by identifying outcomes that would reasonably be associated with public health preparedness and resilience. Specifically, our validation analysis considered confirmed COVID-19 cases, confirmed COVID-19 deaths, influenza vaccination rates, incidence of chlamydia, median household income, and age-adjusted premature mortality. Using both visual comparisons and formal correlation analysis, we were unable to identify a meaningful relationship between the aforementioned metrics and LPHRI indicators or domains. Moreover, some of the metrics showed no correlation with the aggregate domain scores; however, when disaggregated by state or county population size, the correlations differed, suggesting potentially meaningful relationships with state or county size and health outcomes that are not measured by LPHRI.

Strengths and Limitations

In addition to the novelty and strengths of this analysis, there are some limitations that should be noted regarding the data used in the LPHRI. These limitations are all within the normal scope of most research projects and should help guide the user in interpreting and using the data.

While we have made attempts to minimize missing data, there are still instances where data were not readily available. For some indicators, we averaged over two years of data, such that if one of those two years were missing, then the indicator would simply reflect the single non-missing year of data. Additionally, we have employed statistical techniques to impute missing data, which improves the validity of our analytic results.

We recognize that the efforts to modernize public health data systems are essential components of the PHIG and other related initiatives. The National Association of County and City Health Officials' (NACCHO) 2024 profile report on Public Health Informatics identified that almost 3 in 5 local health departments had existing data modernization efforts underway, and the report provides details on the types of informatics efforts ongoing [8]. However, implementation of these initiatives is still in the early stages, so systematic data at the LHJ level are not widely available. As these efforts mature, , even more robust data may be added to the LPHRI in the future. Table 2 includes descriptions of some additional data limitations pertinent to specific indicator groupings.

Table 2. Data limitations pertaining to specific indicators or domains.

Indicators	Limitations
Public health workforce indicators	Job groupings are based on our categorization of job titles. Some agencies have job titles that are less descriptive (e.g. Public Health Associate), and therefore could not be categorized. For this reason, it is possible that we are undercounting the staffing levels at some agencies. Additionally, while some agency data were available by department (e.g. Public Health Department), other agencies did not provide department information; as such, it is possible we may have included some employees who work in other departments (e.g. nurses can work in health services or public health departments), and some agencies have combined "health and human services" departments. Therefore, the ability to distinguish public health employees from other employees within the agency may be limited. Additionally, this method in enumerating public health employees is not able to characterize the effectiveness (output) of the employees.

Clinical laboratory technicians and technologists, Phlebotomists	For non-metropolitan areas, the Bureau of Labor and Statistics only provides regional estimates (e.g. several counties combined). Therefore, many adjacent rural counties will have the same estimate because a regional estimate has been applied.
Public health expenditures	Agencies may have different ways of defining what expenditures are within public health, environmental health, or other departments or categories, so there may be some inherent differences across agencies based on their definitions. The fiscal years used in the LPHRI includes funding to health departments that was provided as part of the COVID-19 public health emergency. Therefore, decreases in public health expenditures following the end of the COVID-19 emergency may simply represent the end of this emergency funding rather than a reduction in core public health funding. Tracking public health expenditures into future years should provide a clearer picture of local public health investments over time. Expenditures data were not adjusted for inflation.
Wastewater surveillance indicators	We applied some simplifying assumptions for sewersheds that serve households in more than one county. Across the 4 states, there are approximately a dozen sewersheds that served 2 counties, and no sewershed served 3 or more counties. For sewersheds that crossed county borders, we assigned these sewersheds to the county where the majority of the sewershed was located. These assumptions lead to underestimating population coverage in the counties that included only a smaller portion of the sewershed, and overestimating the population coverage in counties that included the larger portion of the sewershed.

Given the findings from our validation analysis, where we did not find any evidence of a meaningful association between LPHRI domains or indicators and the aforementioned validation outcomes, we recommend that the LPHRI should be used as a comparative tool to assess relative resource levels and not as an assessment or predictor of the ability for a region to prevent or withstand a future outbreak.

One notable strength of the LPHRI is its focus on local public health preparedness, with unique indicators developed specifically for this project. To our knowledge, there is no publicly available unified dataset of local public health workforce statistics or local public health expenditures. The LPHRI allows users to make comparisons across counties and local health districts and access these unique data elements that can help inform decision-making about preparedness gaps and local investments.

References

1. Leider, J.P., et al., *The Exodus Of State And Local Public Health Employees: Separations Started Before And Continued Throughout COVID-19.* Health Aff (Millwood), 2023. **42**(3): p. 338-348.

2. Leider, J.P., et al., *The State of the US Public Health Workforce: Ongoing Challenges and Future Directions.* Annual Review of Public Health, 2023. **44**(1): p. 323-341.
3. Centers for Disease Control and Prevention. *Data Modernization Initiative*. 2024 April 23, 2024 [cited 2024 October 16]; Available from: https://www.cdc.gov/surveillance/data-modernization/index.html.
4. Centers for Disease Control and Prevention. *Public Health Infrastructure Grant*. 2024 [cited 2024 October 16]; Available from: https://www.cdc.gov/infrastructure-phig/about/index.html.
5. Mays, G.P. and S.A. Smith, *Evidence links increases in public health spending to declines in preventable deaths.* Health Aff (Millwood), 2011. **30**(8): p. 1585-93.
6. Garcia, E.C., et al., *The Clinical Laboratory Workforce: Understanding the Challenges to Meeting Current and Future Needs*. 2021, American Society for Clinical Pathology and Center for Heatlh Workforce Studies, University of Washington: Washington, DC and Seattle, WA.
7. McClary-Gutierrez, J.S., et al., *SARS-CoV-2 Wastewater Surveillance for Public Health Action.* Emerg Infect Dis, 2021. **27**(9): p. 1-8.
8. Patel, K., et al., *Public Health Informatics Profile*. 2024, National Association of County and City Health Officials: Washington, DC.