Determining Community Health Status Priorities in an Online Analytic Processing (OLAP) Environment

James Studnicki[1], John W. Fisher[1]

1 University of North Carolina, Charlotte, College of Health and Human Services, Department of Public Health Sciences

Abstract

Introduction: The determination of priorities is an essential component of community health status assessment. Yet, there is an acknowledged need for a systematic method which will utilize data in standardized comparisons to yield priorities based on objective analyses.

Method: We have deployed a web-based system with: a flexible online analytic processing (OLAP) interface; multiple sources of event-level data conformed to common definitions in a data warehouse structure; and, centralized technical infrastructure with distributed analytical capabilities. The PRIORITIZATION TOOL integrated into the system takes full advantage of the granularity of multidimensional sources of data to: apply a series of defined objective criteria; vary the weight of those criteria and detect the reordering of the rankings in real-time; and, apply the prioritization algorithm to different categories of health status outcomes.

Results: In our example, mortality outcomes for Miami-Dade County, Florida, were considered with three different weighting combinations of the four primary ranking criteria. The resultant analyses return markedly different mortality priority rankings based upon the selection and weighting of the criteria.

Conclusion: Rankings of community health outcomes based on a static set of criteria with fixed weighting factors may not provide sufficient information necessary for priority setting and may, in fact, be misleading.

MeSH Keywords: Public Health Information, Public Health Practice, Community Health Status Assessment, Community Health Priorities.

Correspondence: jstudnic@uncc.edu

Introduction

Community health status assessment is now widely practiced by not-for-profit hospitals seeking to justify their tax exempt status, by community applicants for many federal programs, and by a majority of the nations 2,500 local health departments.[1] An important component of these assessments is some determination of the subject communities' health status priorities. Priority setting should enable the allocation of resources among competing programs, groups or individuals.[2] However, there is very little empirical evidence on priority setting in public health, knowledge of how priority-setting decisions are made is still rudimentary, and many health care and public health resources are expended without a clear understanding of priorities.[3,4,5] Unfortunately, there is also evidence that local health officials are more likely to use subjective criteria rather than evidence based objective criteria when deciding on the most important health issues in their communities.[6] Thus, there is an acknowledged need for a pragmatic and systematic method which will utilize objective data in standardized comparisons allowing decision makers to rely more on the use of hard evidence.[7,8,9] Evidence-based approaches should lead to more consistent and efficient decision making and ultimately result in more strategic allocation of public health resources.[10]

In separate deployments in the states of North Carolina (CATCH) and, more recently, Florida (HealthTrac) investigators have established a web-based analytical environment with: a flexible and powerful online analytic processing (OLAP) interface; multiple sources of multidimensional, event-level data fully conformed to common definitions in a data warehouse structure; enabled utilization of publicly available decision support software tools; and distributed analytic capabilities with centralized technical infrastructure.[11] Utilizing the analytical capabilities of the OLAP analytical environment, the challenge was to create a community health status prioritizing system which would meet these functional requirements: 1) it must systematically apply a series of defined objective criteria; 2) it must be able to rank different types of health status outcomes (e.g., mortality rates, hospitalization rates); and 3) it must provide flexibility in the weighting of the evaluation criteria to enable iterative, "what-if" types of analyses.

Methods

The HealthTrac/ CATCH data warehouses provide a unique opportunity to simultaneously parse community health outcomes across the full set of demographic dimensions at analyst-controlled levels of granularity, and then to rank-order them according to analyst-defined criteria. These capabilities are possible because of the following functional characteristics of the system:

- Event-level health outcomes are available in five data sets – births, mortality, hospitalizations, emergency department visits, and the cancer registry.
- Using ICD-10 codes, mortality indicators may be aggregated by ICD chapter and sub-chapter, the Clinical Classification System (CCS), or the NCHS 113 Common Causes of Death. Rankings may be based on crude rates, age-adjusted rates, or Years of Potential Life Lost.
- Similarly, hospitalizations may be aggregated according to Primary Diagnosis ICD-9-CM codes using the CCS disease definitions. Rankings may be based on crude rates, age-adjusted rates, length of stay total days, or total charges, depending on the analyst's definition of impact.

2

- Cancer registry data affords access to cancer incidence data aggregated based on site (e.g. breast, lung) definitions, with rankings based on crude rates, age-adjusted rates, and stage-of-diagnosis measures.
- Potential issues with small cell sizes can be addressed by aggregating across time, diagnosis, and/or selected demographic dimensions.

We assume three nearly orthogonal dimensions compose the concept of "priority": impact; departure from norms; and time. These dimensions are defined and operationalized in the following manner:
- The impact dimension measures the effect of the event on the population. Mortality rates (number of deaths per unit of population) and YPLL, lost years of potential life, are measures of the impact of premature mortality. For hospitalizations, impact measures may include admission rates, total length of stay days, or total charges.
- The second dimension, departure from norms, measures how much the measure differs from some benchmark. Typical standards include state or peer averages, or standard deviations from the mean of all measures in a selected geography (e.g., county or hospital service area).
- A third dimension - temporal rate of change - simply indicates whether an indicator is measurably improving, deteriorating, or not reflecting any significant change trend.

The relative importance ascribed to each of these dimensions can have a major effect on the rankings of the indicators, and this is a major limitation of static conventional ranking methodologies that use county-level comparative rankings as the sole measure. A common problem with many county-level ranking systems is their attempt to aggregate mixed-outcome types into a single "health" index or ranking. Combining rankings representing mortality, hospitalizations, health resource availability, disease morbidity and other determinants of health requires at least an implicit relative weighting of the various outcome types. Assigning an explicit comparative factor to each outcome type is a highly subjective exercise; absent such an assignment, each of these indicators is implicitly assumed to have equal effect on the health of the community.

The HealthTrac/CATCH prioritization tool avoids this conundrum by first dividing the indicators into five separate categories representing the disparate outcome types:
- Mortality – computed from vital statistics death records using primary cause of death;
- Morbidity – selected from county-level reports of infectious and sexually transmitted diseases;
- Hospitalizations – Inpatient (at least overnight admissions) hospitalization rates using the Primary Diagnosis ICD-9-CM code aggregated using the Clinical Classification System hierarchy;
- Behavioral – county-level rates of conditions/behaviors selected from the annual Behavioral Risk Factor Surveillance System (BRFSS) survey;
- Correlates of health (violent crime, high school drop-out rates, uninsured rate, etc) – amalgamation of county-level indicators from a wide variety of public data sets.

The tool then employs a conventional top-down approach using county-level aggregates and four measures representing the three priority dimensions of impact, departure from norms, and time:

- Impact is measured by counting the number of individuals affected. Additional measures warranting investigation include Length of Stay and Total Charges for hospitalizations, and Years of Potential Life Lost (75) for mortality.
- Departure from Norms is calculated as rate differences between the subject county and the rates for both the State and the county's peers. Additional benchmarks available include the national average and any published goals, such as Healthy People 2020.
- The Temporal dimension is captured by measuring any significant trend over a selected period such as the most recent 3 years or 5 years. Trend lines are computed using a linear regression curve and an adjustable minimum R^2 value for significance.

The challenge of normalizing the dimensional variances to allow their aggregation by simple addition after being weighted is addressed by assuming a normal distribution and converting the values to standard deviations (z-values). For instance, for the mortality category, the number of deaths due to each of the top common causes of death vary widely within a given county. By computing the distance in standard deviations from the median number of deaths for all causes of death, the very wide differences are reduced to a much more manageable range and one that is common to the other criteria as well.

Results

Once the prioritization tool is activated, the user will be prompted to select the category of outcomes to be rank ordered. In our example, the user selects the mortality outcomes (Figure1). On the left hand side of the screen, four screening criteria previously defined are listed: number affected; trend; magnitude of difference (peer); and magnitude of difference (state). A fifth criterion, community support, may also be activated. This enables any special influence which may be specific to the community environment to be considered in the prioritization process.

To the right of the criteria are slide bars which serve to determine the weight given to each of the criterion. The analyst locates the slide bar at the exact point to represent the relative contribution of each criterion to the ranking algorithm. The total weight of all criteria combined is exactly 1.

Figure1. Health Outcome Category Selection

In our first example (Figure2), we have weighted the number of persons affected and the direction of the trend in the outcome as about equally important in the ranking. The difference between the values in the subject community (Miami-Dade) versus the values in a sociodemographically equivalent peer group of counties, and the Florida state average values are not considered in this ranking. This weighting combination produces a ranking with chronic and acute lower respiratory disease death rates as the highest mortality priorities.
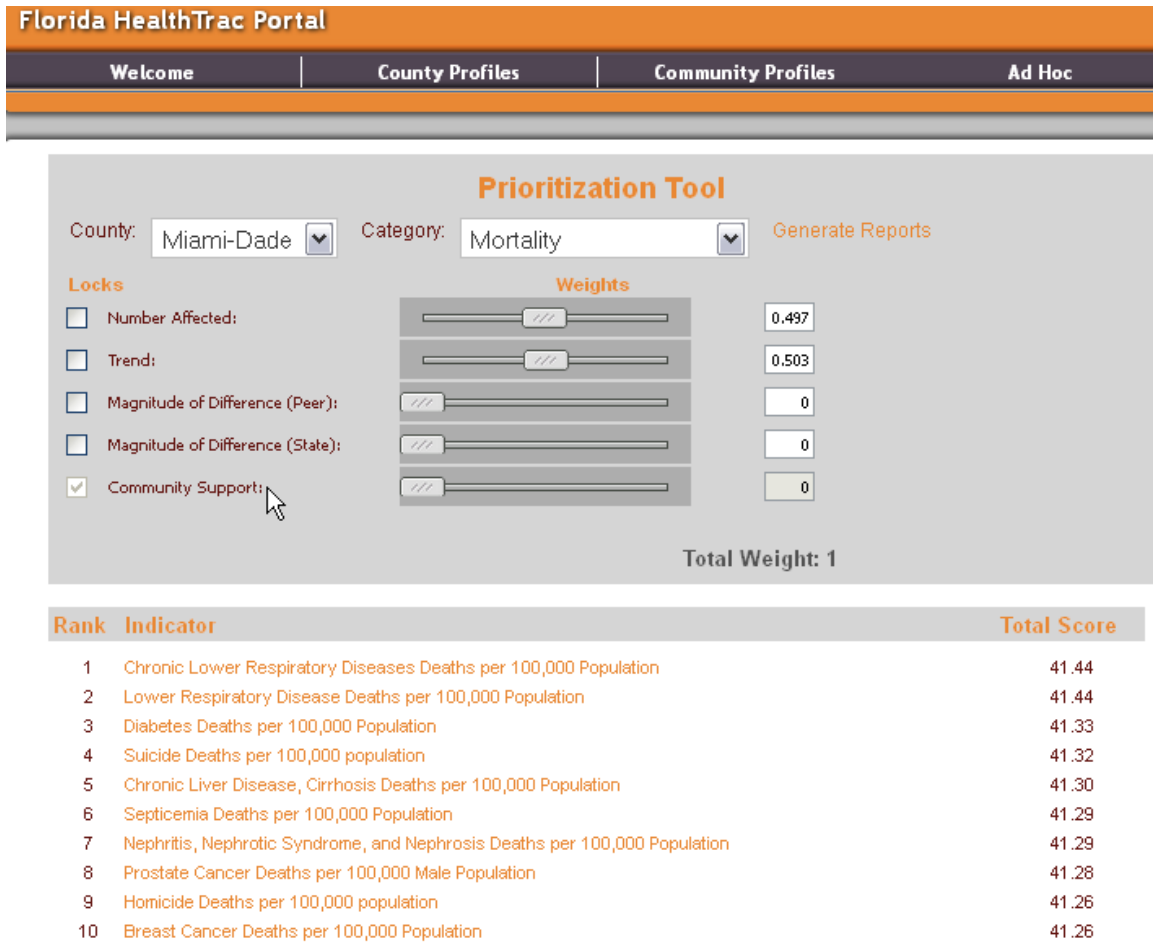
Figure 2. Criteria Weighting Combination No.1

In our next iteration (Figure 3), we eliminate the influence of temporal trend and weight the algorithm to produce a ranking of causes of death solely on the number of persons affected. The ranking of mortality outcomes, now strikingly different, places the two measures of premature death (years of potential life lost 65 and 75) at the top of the rankings because they aggregate deaths from all causes. Note also, that the familiar major causes of death (heart disease, cancer, stroke, cerebrovascular disease) emerge as high priority mortalities in this ranking. Lower respiratory diseases remain in the listing but are no longer top ranked.
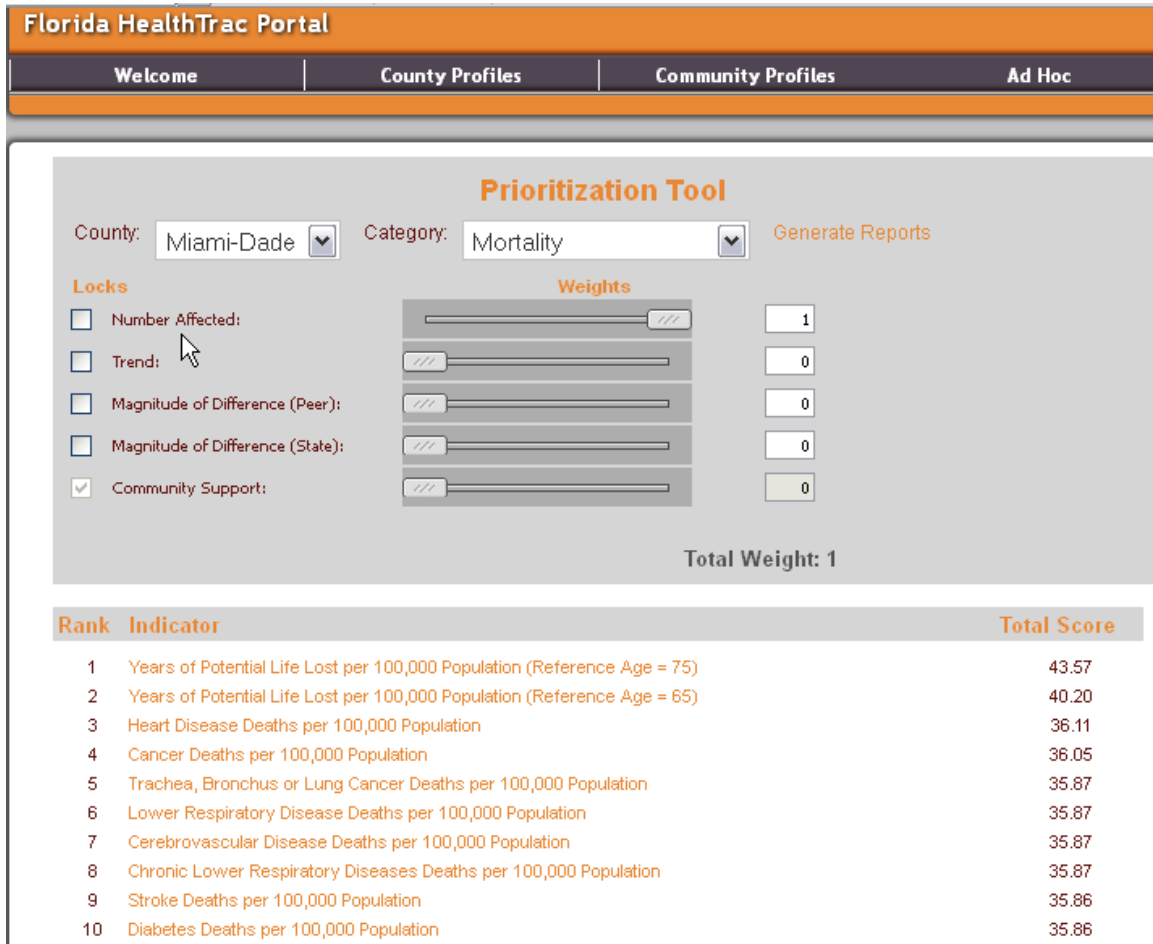
Figure 3. Criteria Weighting Combination No.2

Finally, we allow the rankings to be solely determined by Miami-Dade's mortality profile compared to their peer counties and the state average (Figure4). This time the number affected and temporal trend are not considered. Incidentally, this is the way most of the static community report cards primarily derive their rankings. Again, the rankings are markedly changed with HIV/AIDS, Homicide and Diabetes emerging at the very top of the rankings. Deaths from lower respiratory diseases have disappeared from the top ten.
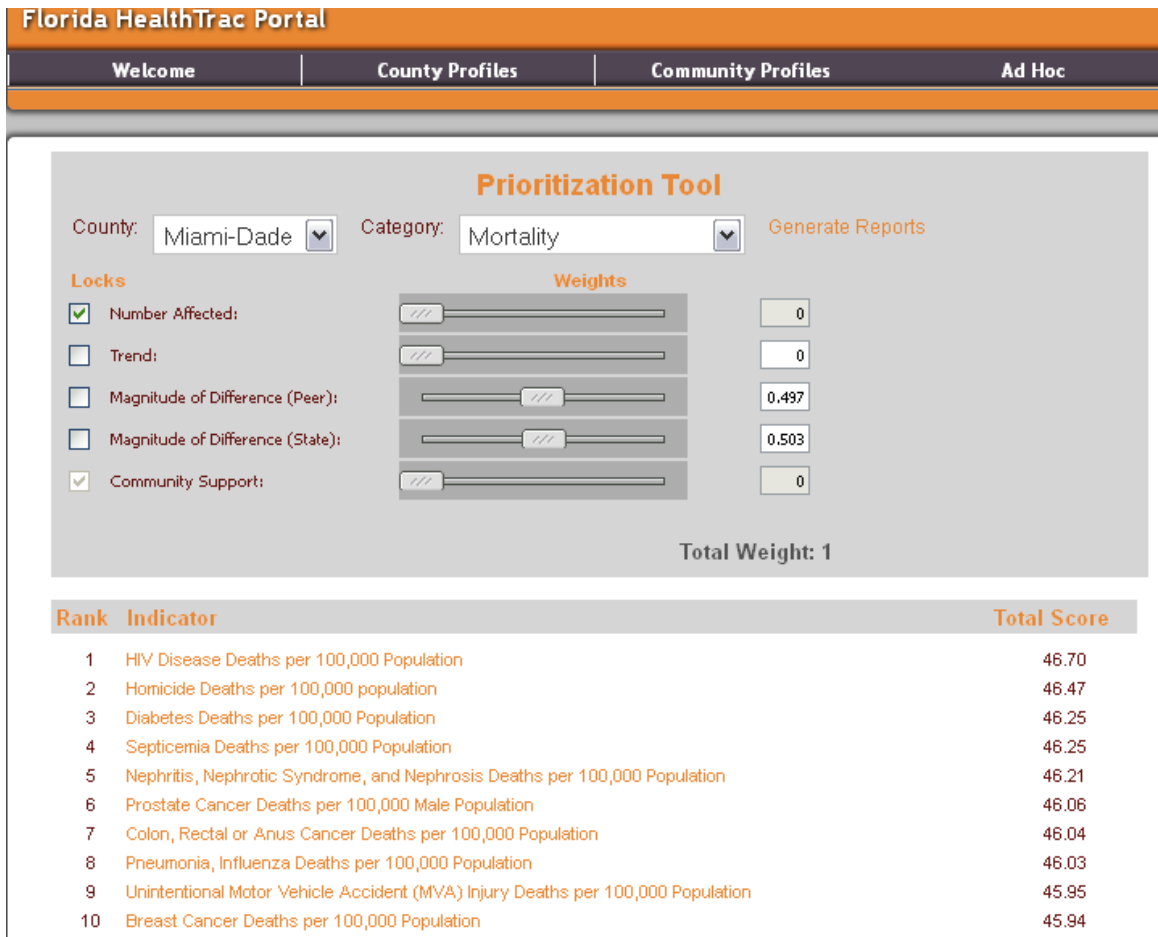
Figure 4. Criteria Weighting Combination No.3

Discussion

Clearly, variation in the explicit weighting of objective criteria can dramatically change the rankings of health outcomes which are used to derive community health status priorities. Priorities based upon a static ranking of a number of fixed indicators are unable to reflect the unique and changing circumstances and characteristics of local populations. The prioritization tool herein described utilizes county level aggregates to identify the high priority issues, and then assumes that the OLAP drill down capabilities will be used to investigate the distribution and patient characteristics of high-priority indicators. In the case of premature death as indicated by years of potential life lost (YPLL), for example, the OLAP tool is available to identify the contribution made by various causes of death (e.g., infant mortality, heart disease, homicides), in specific age groups, by race, in various geographic locations within the county. Thus, the analyst can produce actionable information about the problem. In the absence of an OLAP enabled warehouse of event-level data sources, this macro-micro sequential analysis is not possible.

In the future, OLAP capabilities will allow for a "bottom-up" rather than the current "top-down" approach. The proposed approach starts at the lowest grain of geography using a default weighting scheme to allow the analyst to roll up dimensions (e.g., collapse the race and/or gender

8

dimensions) to measure statistical power. This approach will produce high priority POPULATIONS rather than causes of death or reasons for hospitalization. The populations will be defined by the outcome, but also age, race, location, and other characteristics. Why is this capability necessary in understanding community health status priorities?

Effective health outcome interventions are local; that is, they address specific health issues affecting identifiable sub-groups of the population defined by geography, race, age, and gender. Prioritizing population health outcome issues for interventions requires identifying the outlying health outcomes for these sub populations. As we have seen, this has been accomplished by comparing county-level values for a limited number of generic indicators against other counties, and then "drilling down" to identify variations within the county to identify the potential out-lying sub populations. This approach works best for small-to-mid-size, fairly homogenous counties. For large, heterogeneous counties, the extreme outliers will tend to "wash out" each other, and the county level average value may not in any way reflect the true depth of the problem for the most vulnerable segments of the population. Consider that for Florida, there are some 1,000 zip codes. Each zip code population can be subdivided by race (3 values - Black, White, and Other), 2 genders (Male, Female), and perhaps 5 age bands. This yields 30,000 sub population groupings. Using the NCHS 113 common causes of death, the mortality possibilities alone total over 3.3 million possible " sub-population "problem outcomes (113 x 30,000). Searching this space manually is simply infeasible. Yet, the event-level data sources and information technology is available to deliver this type of analytical capability to every US community.

Acknowledgment

Corresponding Author

Jams studnicki, Sc.D.
Irwine Belk Endowed Chair and Professor[1]
jstudnic@uncc.edu
Phone: 704-687-8981
Fax: 704-687-6122

References

[1] National Association of County and City Health Officials. 2005 National Profile of Local Health Departments. Washington, DC, 2006.

[2] McKneally M, Dickens B, Meslin E, et.al. Bioethics for clinicians: resource allocation. Canadian Medical Association Journal. 1997;157(2):163-167.

[3] Martin D, Singer P. A strategy to improve priority setting in health care institutions. Health Care Analysis. 2003;11(1):59-68.

[4] Martin D, Abelson J, Singer P. Participation in health care priority-setting through the eyes of the participants. Journal of Health Services Research and Policy. 2002;7(4):222-229.

9

[5] Woolf S, Stange K. A sense of priorities for the healthcare commons. American Journal of Preventive Medicine. 2006;31(10):99-102.

[6] Platonova E, Studnicki J, Fisher J. Local health department priority setting: an exploratory study. Journal of Public Health Management and Practice. 2010;16(2):140-147.

[7] Mitton C, Donaldson C. Tools of the trade: a comparative analysis of approaches to priority setting in healthcare. Health Services Management Research. 2003;16(2):96-105.

[8] Mitton C, Patten S, Waldner H, et.al. Priority setting in health authorities: a novel approach to a historical activity. Social Science and Medicine. 2003;57:1653-1663.

[9] Vilnius D, Dandoy S. A priority rating system for public health programs. Public Health Reports. 1990:105(5):463-470.

[10] Sibbald S, Singer P, Upshur R, et.al. Priority setting: what constitutes success? A conceptual framework for successful priority setting. BMC Health Services Research. 2009;4:43.doi:10.1186/1472-6963-9-43.

[11] Studnicki J, Fisher J, Eichelberger C, et.al. NC CATCH: advancing public health analytics. Online Journal of Public Health Informatics. 2010; 2(3).

10