What Makes a Good Measure?

Michael A. Stoto, PhD
Georgetown University and
Harvard School of Public Health
June 6, 2008
Maple Hill Farm, Hallowell ME

Learning Objectives

- Construct well-operationalized measures
- Appreciate statistical issues with rates and proportions
- Understand how to balance validity, reliability, and responsiveness to change
- Protect privacy and confidentiality

Outline

- ***Principles of community health assessment and performance measurement***
- Statistics of rates and proportions
- Balancing validity, reliability, and responsiveness to change
- Exercise
- Privacy and confidentiality

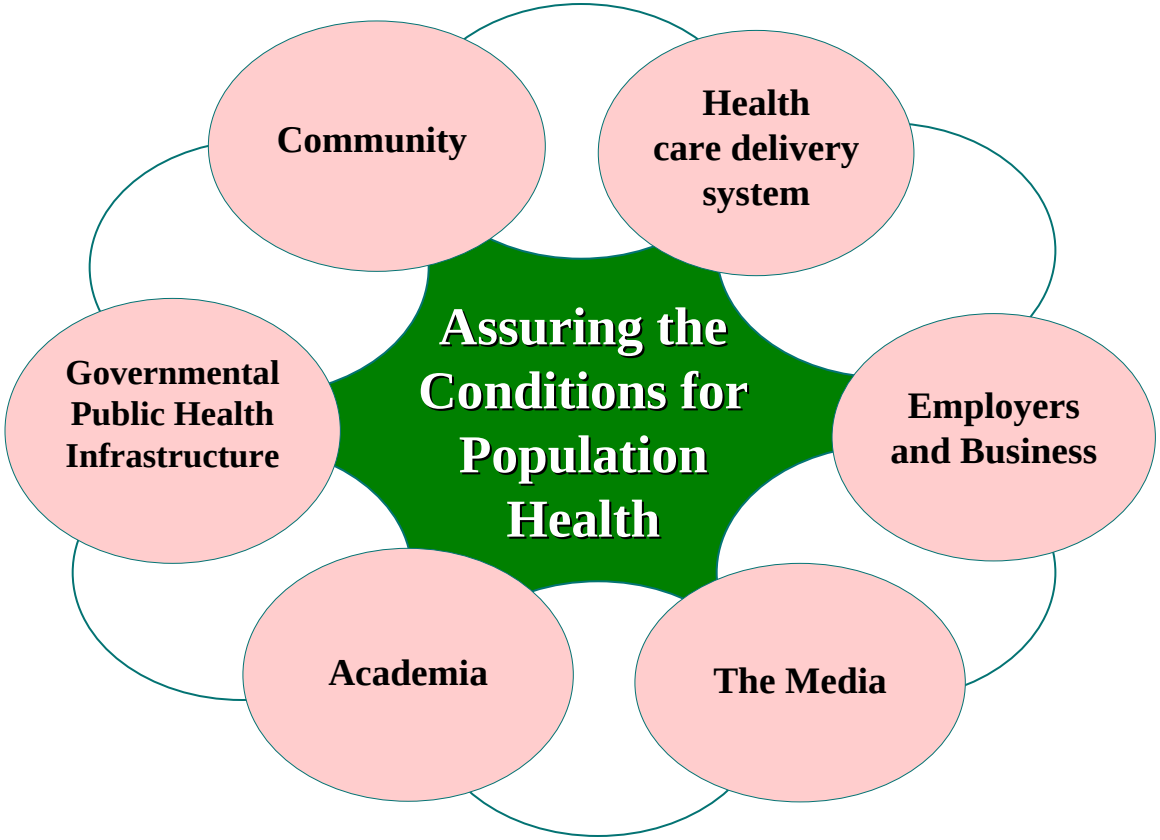
“Population Health”

- Kindig and Stoddart: The health outcomes of a group of individuals, including the distribution of such outcomes within the group
- Characteristics of the population health perspective
  - Broader array of the determinants of health than in traditional public health
  - Recognition that responsibility for health is diffuse
- Management of population health
  - Systems perspective, partnerships
  - Increasing amounts of public health activities are managed not within but between institutions

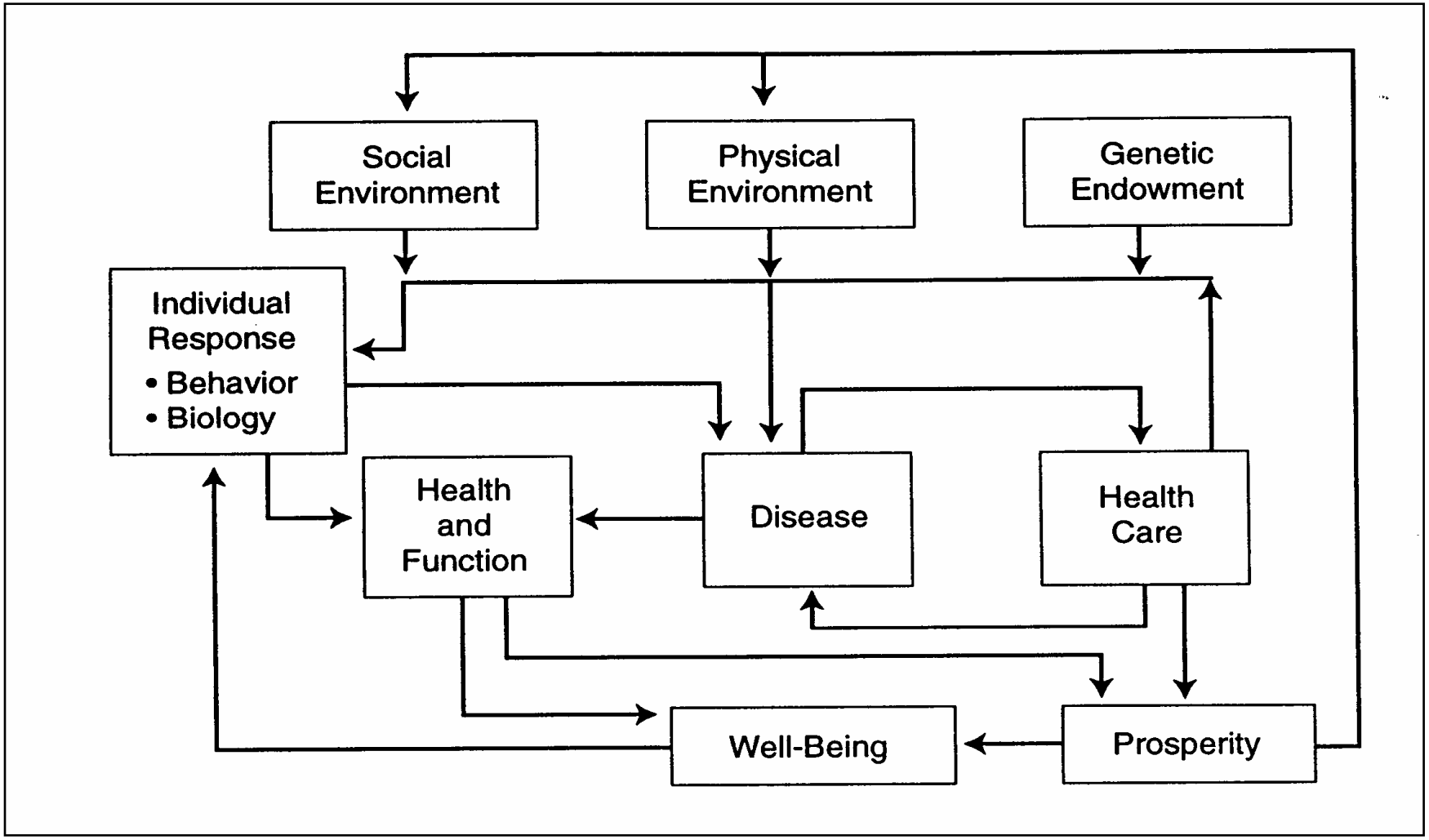
IOM: *The Future of the Public's Health in the 21st Century*

- Public health: What society does collectively to assure the conditions for people to be healthy
- Multiple determinants of health
  - Health of populations and individuals is shaped by a wide range of factors in the social, economic, natural, built, and political environments
  - Health care services and biomedical technologies can generally address only the immediate causes of disease
  - Healthy policies address education, adequate housing, a living wage, clean air, socio-economic inequalities, etc.

The Public Health System



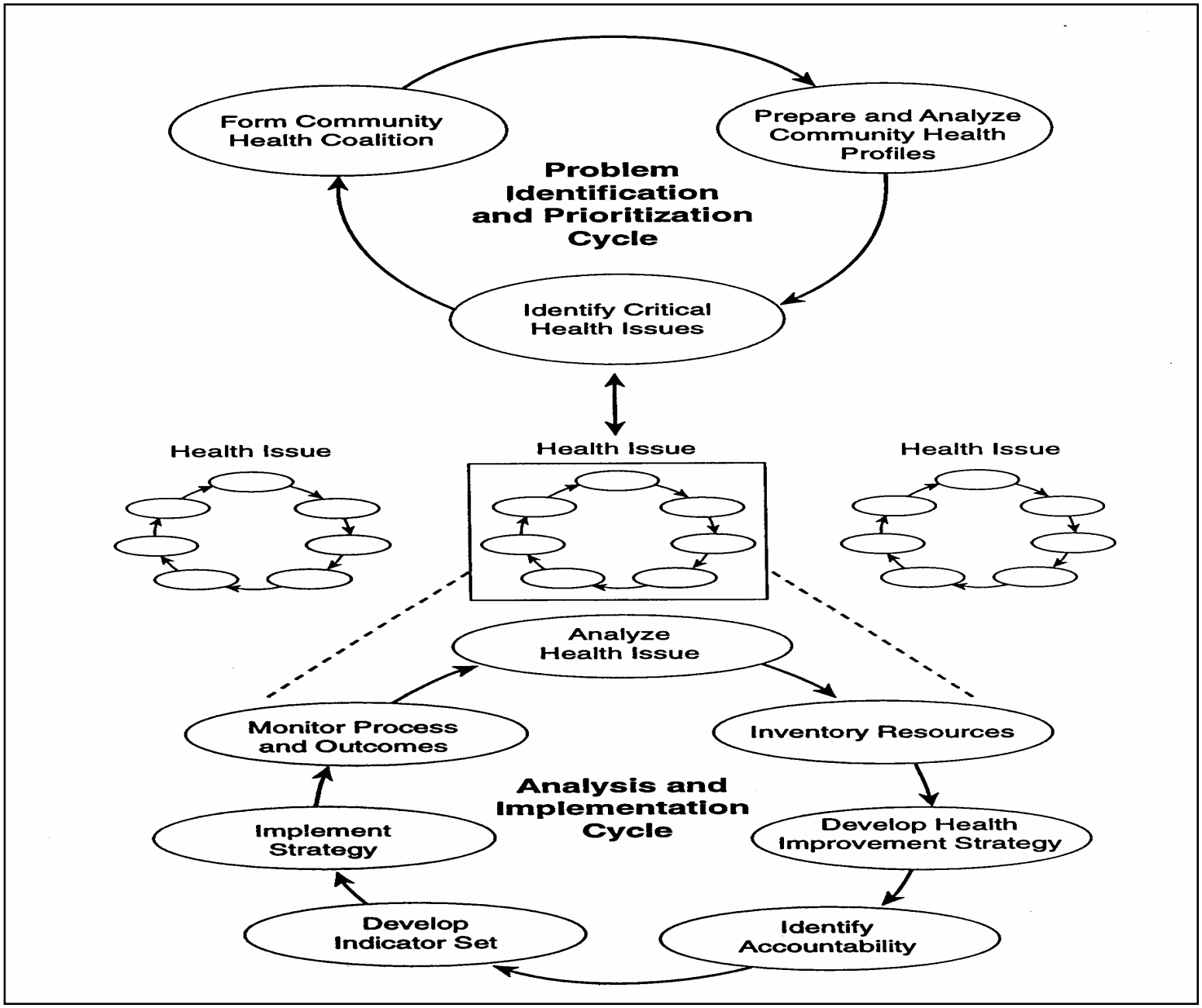
Evans and Stoddart "Field Model"



IOM Community Health Improvement Process (CHIP)

Improving Health in the Community, IOM, 1997

- **The community's health depends on the interaction of many factors, entities, organizations, and interests in the community**
- **Community health is a** ***shared responsibility***
  - **Community health assessment**
- **Specific entities in the community must be** ***accountable*** **for the actions that they can take to improve community health**
  - **Entity-specific performance measures**


Form Community Health Coalition
Prepare and Analyze Community Health Profiles
Problem Identification and Prioritization Cycle
Identify Critical Health Issues
Health Issue
Health Issue
Health Issue
Analyze Health Issue
Monitor Process and Outcomes
Inventory Resources
Analysis and Implementation Cycle
Implement Strategy
Develop Health Improvement Strategy
Develop Indicator Set
Identify Accountability

MAPP: Mobilizing for Action through Planning and Partnerships

Summary – session 1

- The community's health depends on the interaction of many factors, entities, and organizations in the community
- Community health profiles/assessment
  - summarize the health of a community
  - measure our *shared responsibility*
- Key characteristics
  - a *limited* yet *comprehensive* set
  - of *coherent* and *significant* indicators
  - that can be *monitored* over time and
  - *disaggregated* to relevant social units

Summary – session 2

- Performance measures hold specific entities accountable for actions that they can take to improve community health
- Key characteristics
  - Accountable entity identified
  - Balance among structure, process, and outcome measures
  - Evidence-based link between performance and health
  - Validity, reliability, and sensitivity to change

Outline

- Principles of community health assessment and performance measurement
- ***Statistics of rates and proportions***
- Balancing validity, reliability, and responsiveness to change
- Exercise
- Privacy and confidentiality

Measurement theory & methods

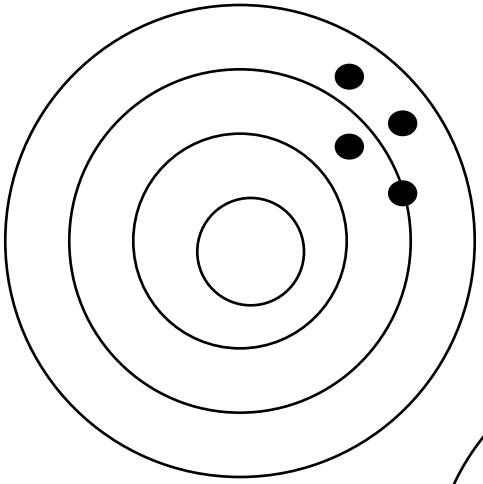
Steps for developing measures

1. Clarify the purpose of measurement
2. Identify the concepts to be measured
3. Identify specific indicators of these concepts
4. ***Assess validity, reliability, responsiveness to change***

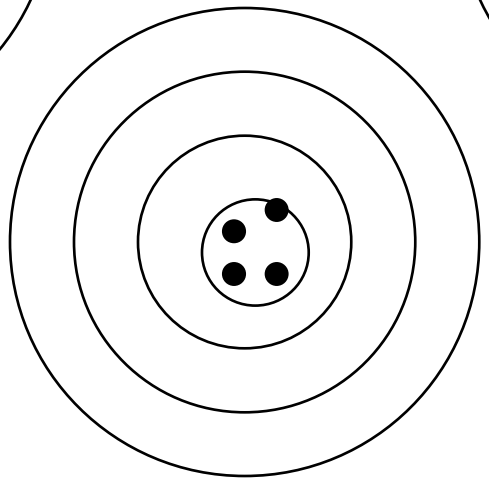
Assess the proposed measures

- Reliability
  - is the indicator **consistently** measuring the concept?
  - Is measurement error small compared to population variability?
- Validity
  - is the indicator measuring the **right concept**?
- Robustness and responsiveness to change
  - will the indicator change if and only if the concept being measured changes?

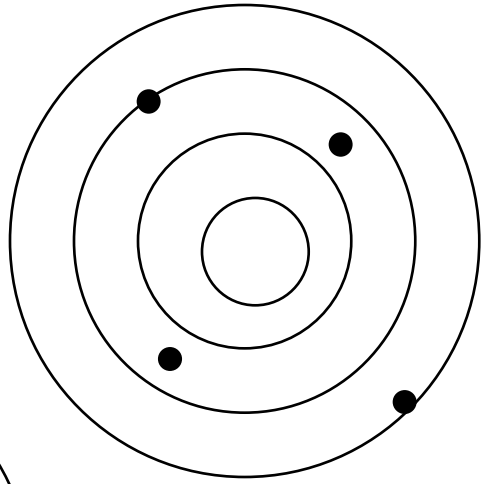
Reliability and validity



Reliable



Reliable and valid



Valid

Statistical reliability due to small numbers

- Data systems based on case reports
  - *x* forms numerator of a rate *p*
    - base population = *n*
    - E.g. prevalence of condition A per 1,000 or 100,000
    - Stochastic variability of the disease process generates 3, 8, 6, 7 cases per year with E(*x*) = 5
  - Statistical models:
    - binomial, Poisson, etc.

Statistical reliability due to small numbers

- Data systems based on case reports
  - x forms numerator of a rate p
    - base population = n
  - When p is small
    - Var (x) = np
    - Var (p) = Var (x/n) = p/n
  - If expected number of cases (np) = 4
    - Std. Dev. = $\sqrt{4} = 2$
    - 50% increase from 4 to 6 = 1 std. dev.

Statistical reliability due to small numbers

- Data systems based on case reports
  - *x* forms numerator of a rate *p,* population = *n*
    - When p is small Var (*p*) = Var (*x*/*n*) = *p*/*n*
  - Denominator is also important

x = 4, n = 100 or 10,000		
x = 4	n = 100	SD(p) = √0.04/100 = 0.02
x = 4	n = 10,000	SD(p) = √0.0004/10,000 = 0.0002
p = 0.04, n = 100 or 10,000		
p = 0.04	n = 100	SD(p) = √0.04/100 = 0.02
p = 0.04	n = 10,000	SD(p) = √0.04/10,000 = 0.002

Statistical reliability due to small numbers

- Sample data
  - With simple random sampling
    - Var (p) = $p(1-p)/n$
    - Small sub-sample may be unbalanced with respect to age, race, sex, etc.
  - For complex sampling designs formula is more complicated, but depends similarly on n
    - Some small areas may have no sample elements

Statistical reliability due to small numbers

- Conclusion: variability of public health data is generally inversely proportional to *n*
  - For epidemiologic and demographic rates, *n* is the size of the population generating events
    - *n* cannot be changed
  - For survey data, *n* is the sample size in the community (not population size *N)*
    - Same *n* needed for communities with 5,000 or 5 million inhabitants

Addressing statistical reliability due to small numbers

- Increase sample size
  - Not possible for record based data
  - Possible but expensive for sample data
- Aggregate data
  - from 3 to 5 previous years
    - May mask changes over time
    - Results are “out of date”
    - For small communities, still may not be enough
  - from adjacent communities
    - May hide meaningful differences among those communities
  - for racial/ethnic groups
    - May mask important health disparities

Addressing statistical reliability due to small numbers

- Suppress results (counts, rates, averages) based on less than *x* observations
  - "Rule of 5" or "Rule of 3"
- Report only counts, not rates or averages
  - Rationale: remind user of lack of precision
  - Users can calculate rate themselves
- Calculate confidence intervals
  - Optional or automatic
  - Interpretation
    - Sampling theory for samples
    - Stochastic process for case reports
- Statistical tests
  - χ^2 test for trend over time
  - χ^2 test for difference in rates between groups

Outline

- Principles of community health assessment and performance measurement
- Statistics of rates and proportions
- ***Balancing validity, reliability, and responsiveness to change***
- Exercise
- Privacy and confidentiality

Validity

- Face validity: do the indicators measure the concept we want to measure?
- Real vs. artifactual differences
  - changes in mortality statistics, self-reported health conditions, etc.
  - changes in the social process of measuring health
  - age-standardization, risk adjustment

Concepts and specific indicators for state Maternal and Child Health Programs

Concept	***Specific indicator***
Insurance coverage	% of children with special health care needs age 0 to 18 whose families have adequate private and/or public insurance to pay for the services they need.
Insurance coverage	% of children without health insurance.
Adequate prenatal care	% of infants born to pregnant women receiving prenatal care beginning in the first trimester.
Immunization coverage	% of 19 to 35 month olds who have received the complete immunization schedule for diphtheria-tetanus-pertussis, polio, measles-mumps-rubella, *H. influenza*, and hepatitis B before their second birthday.

HP2010 Leading Health Indicators and possible local data

Concept	***Specific indicator***
Physical activity	% engaging in physical activity (BRFSS)
	Coronary heart disease death rate
Overweight and obesity	% overweight or obese (BRFSS)
	Diabetes death rate
Responsible sexual behavior	Gonorrhea incidence
	Birth rate for girls aged 15-17
Mental health	Suicide rate
	Mental health "not good" days (BRFSS)
Injury and violence	Motor vehicle crash death rate
	Firearm-related death rate
	Reported rape or attempted rape
Environmental quality	Days ozone standards exceeded
	Cases of foodborne pathogens (E. Coli 0157:H7, Salmonella)

National Diabetes Quality Improvement Alliance *Performance measurement set for adult diabetes*

Concept	***Measure*** (per year, denominator = all diabetes patients aged 18-75)
Intensive therapy of hemoglobin A1c	• % of patients with ≥ 1 A1c test • % of patients with most recent A1c > 9%
Lowering serum cholesterol	• % of patients with ≥ 1 LDL-C test • % of patients with LDL-C < 130 mg/dl
Early detection of ESRD	• % of patients with ≥ 1 microalbumin test
Early detection of retinopathy	• % of patients with standard eye exam (ok in previous year if low risk)
Prevention of foot ulcers/amputations	• % of patients with foot exam

Reliability vs. validity

- Use running averages and statistical smoothing
  - example: average infant mortality rate over 3 or 5 years
  - trade timeliness for reliability
- Use proxy measures that reflect trends and differences
  - example: low birth weight rather than infant mortality
  - trade validity for reliability

Example: Promoting successful birth outcomes

- Example based on *Access to Health Care in America* (IOM, 1993)
- One of five objectives: Promoting successful birth outcomes
- Combination of 4 utilization and outcome indicators

Example: Promoting successful birth outcomes

- Indicator #1 (utilization measure)
  - Percentage of pregnant women obtaining adequate prenatal care
  - Strength: direct measure of access
  - Weaknesses:
    - Measured by initiation and frequency
      - neither alone sufficiently measures adequacy
        - » nothing on when visits occur, content, continuity
        - » large number could represent a problem
    - more complex indices can be confounded by missing or incomplete data
    - recall problems

Example: Promoting successful birth outcomes

- Indicator #2 (outcome measure)
  - Infant mortality rate
  - Strengths
    - commonly used measure of access
    - available everywhere from vital statistics
  - Weaknesses
    - provides little information about access barriers
    - rate includes causes of death that cannot be affected by the health care system
    - high variability in infant deaths in some areas

Example: Promoting successful birth outcomes

- Indicator #3 (outcome measure)
  - Low birthweight rate
    - proportion of infants weighing less than 2500g
  - Specific to adequate prenatal care and access to nutrition services
  - Strengths
    - important predictor of infant survival
    - numerator not as rare, so more stable
  - Weakness
    - timeliness of published data

Example: Promoting successful birth outcomes

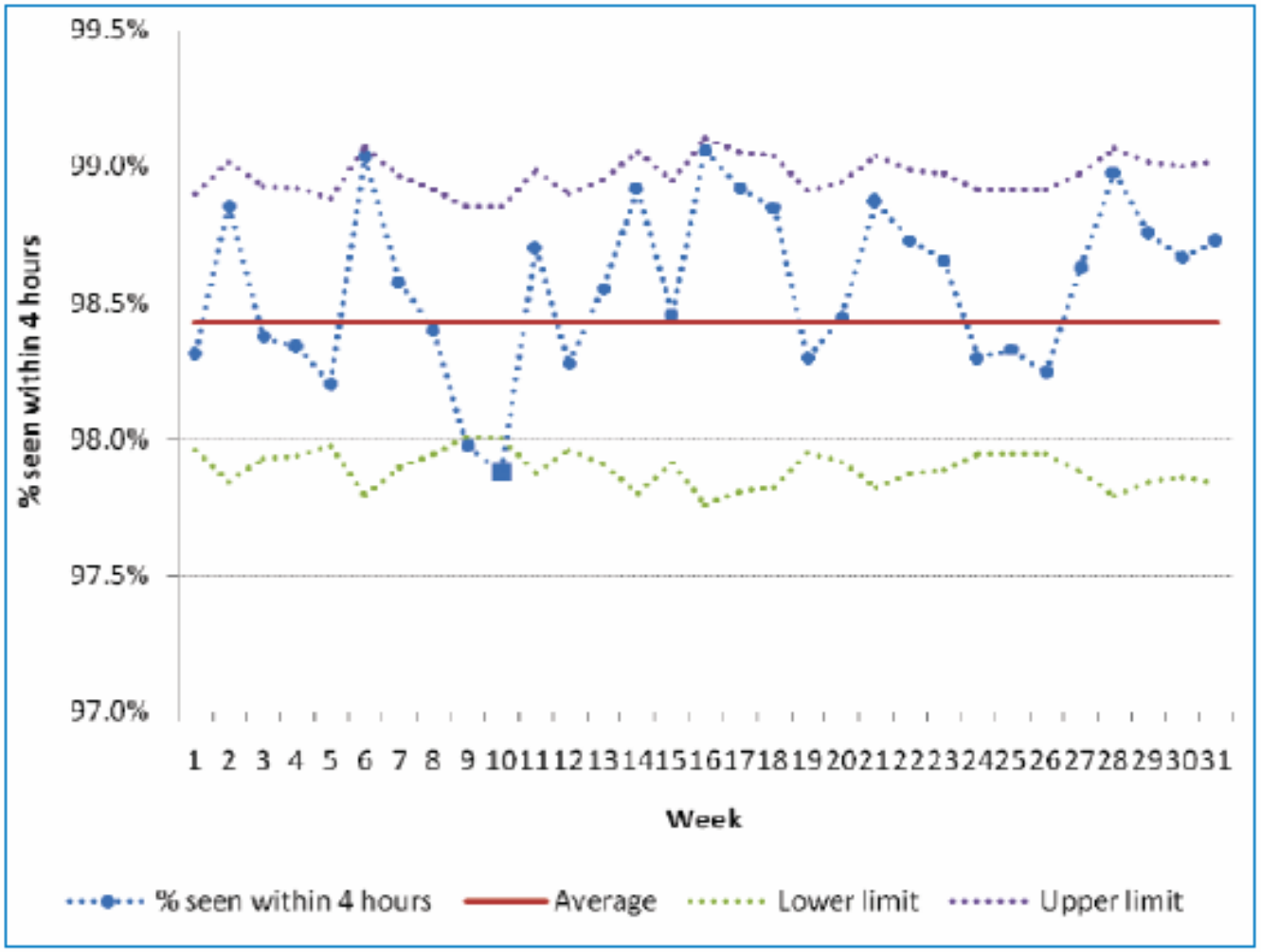
- Indicator #4 (outcome measure)
  - Congenital syphilis rate
  - Strength
    - reportable condition in most states
    - very specific to lack of or inadequate prenatal care
  - Weaknesses
    - reporting may be incomplete
    - syphilis is rare in most states

Responsiveness to change: understanding variation

- Types of variation
  - Common cause: normal, everyday, inevitable (& usually unimportant) that is intrinsic and natural
  - Special cause: indicative of something special happening that calls for fuller understanding and often action
  - Temporal vs. cross institution or area
- Statistical Process Control (SPC)
  - Valid approach to distinguish between the two
  - Persuasive way to present data to others

Figure 2: A statistically uniform system (compared with itself over time)

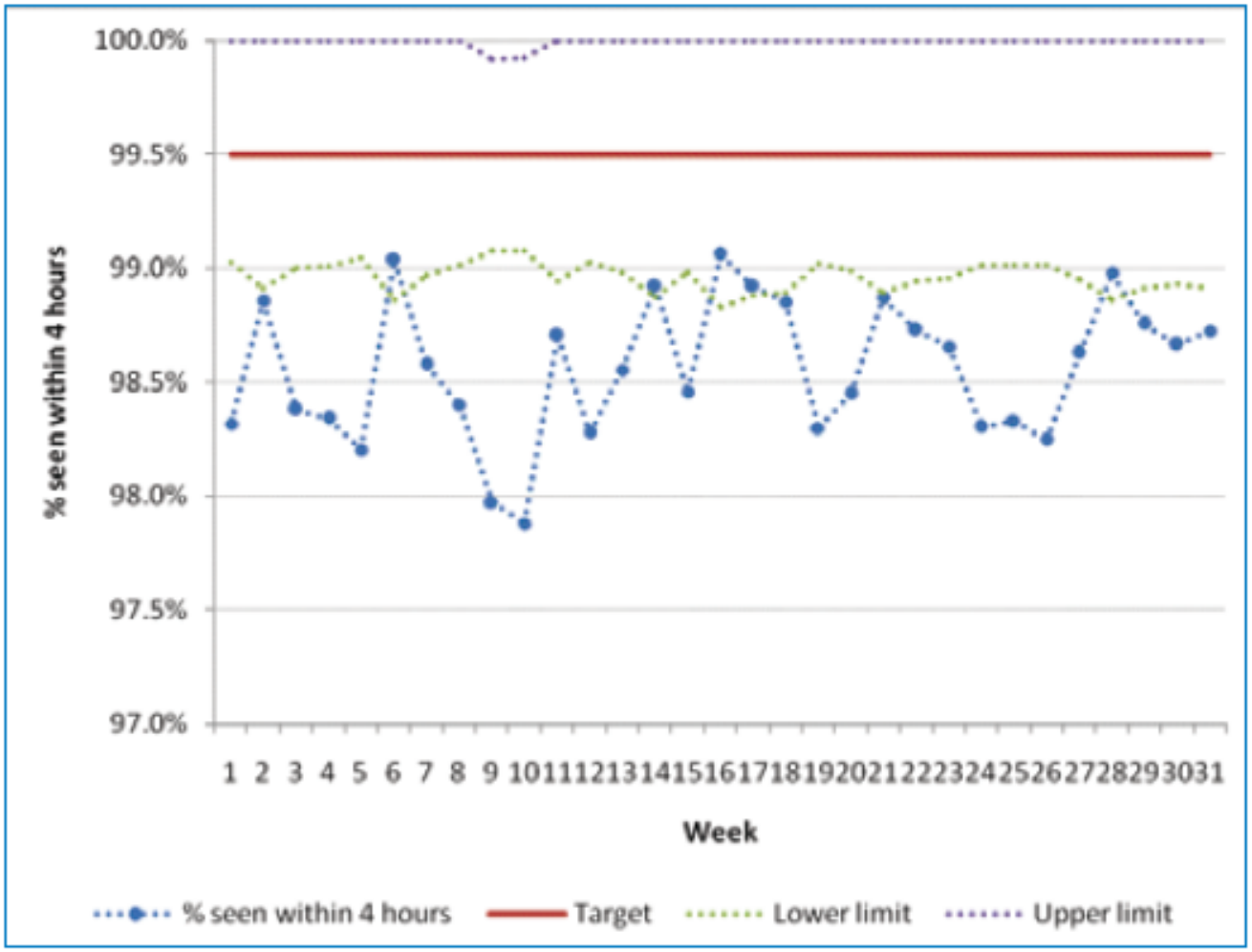
Showing the proportion of A&E attendances seen within four hours. The average proportion is 98.4 per cent - there is one special cause variation denoted by the square marker at week 10.



Source: NHS (UK), The Good Indicators Guide

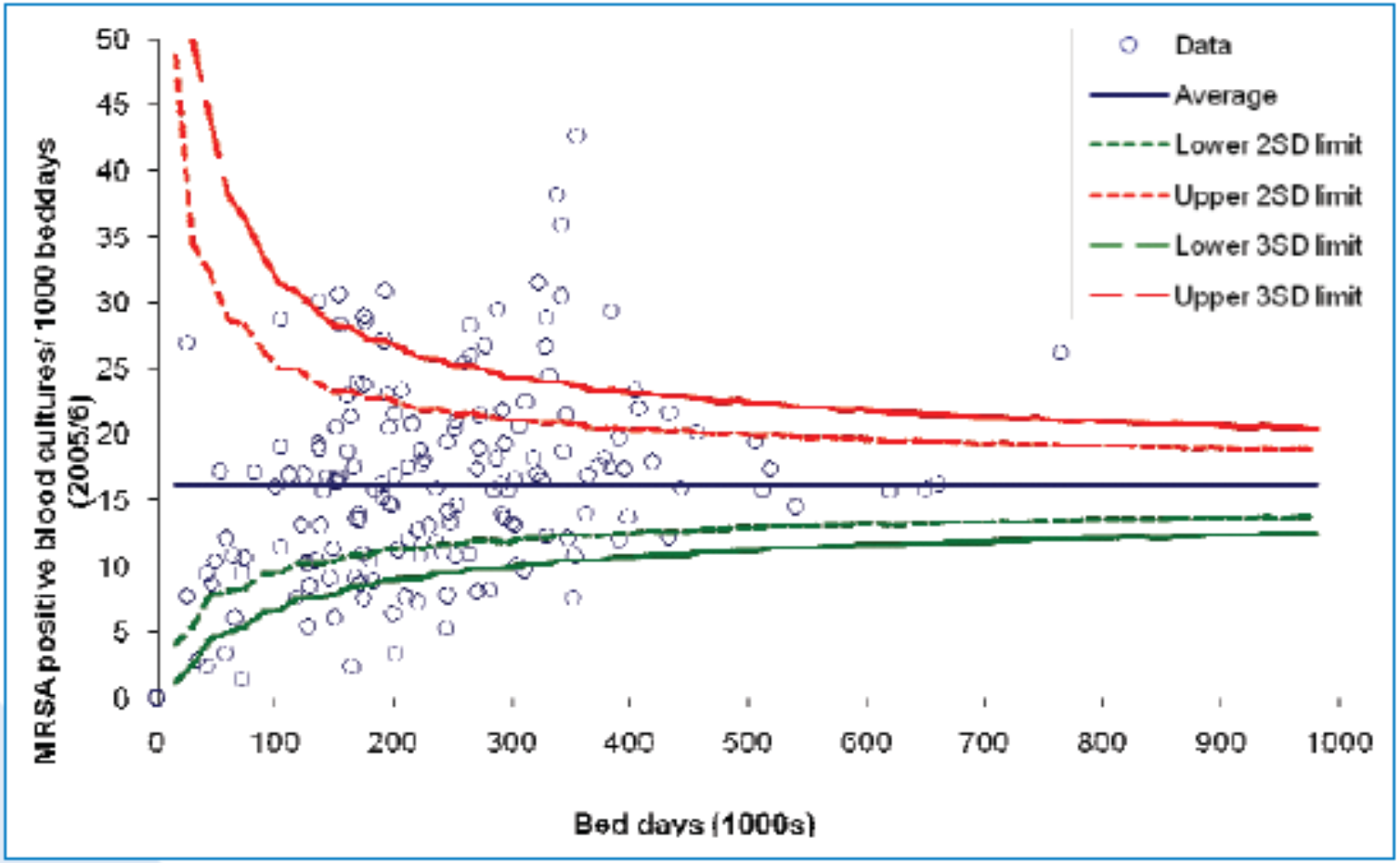
Figure 3: A statistically uniform system not meeting expectations [compared with itself over time]

Showing the proportion of A&E attendances seen with four hours. However, compared with target expectation (red line), the whole system is performing at the wrong level.



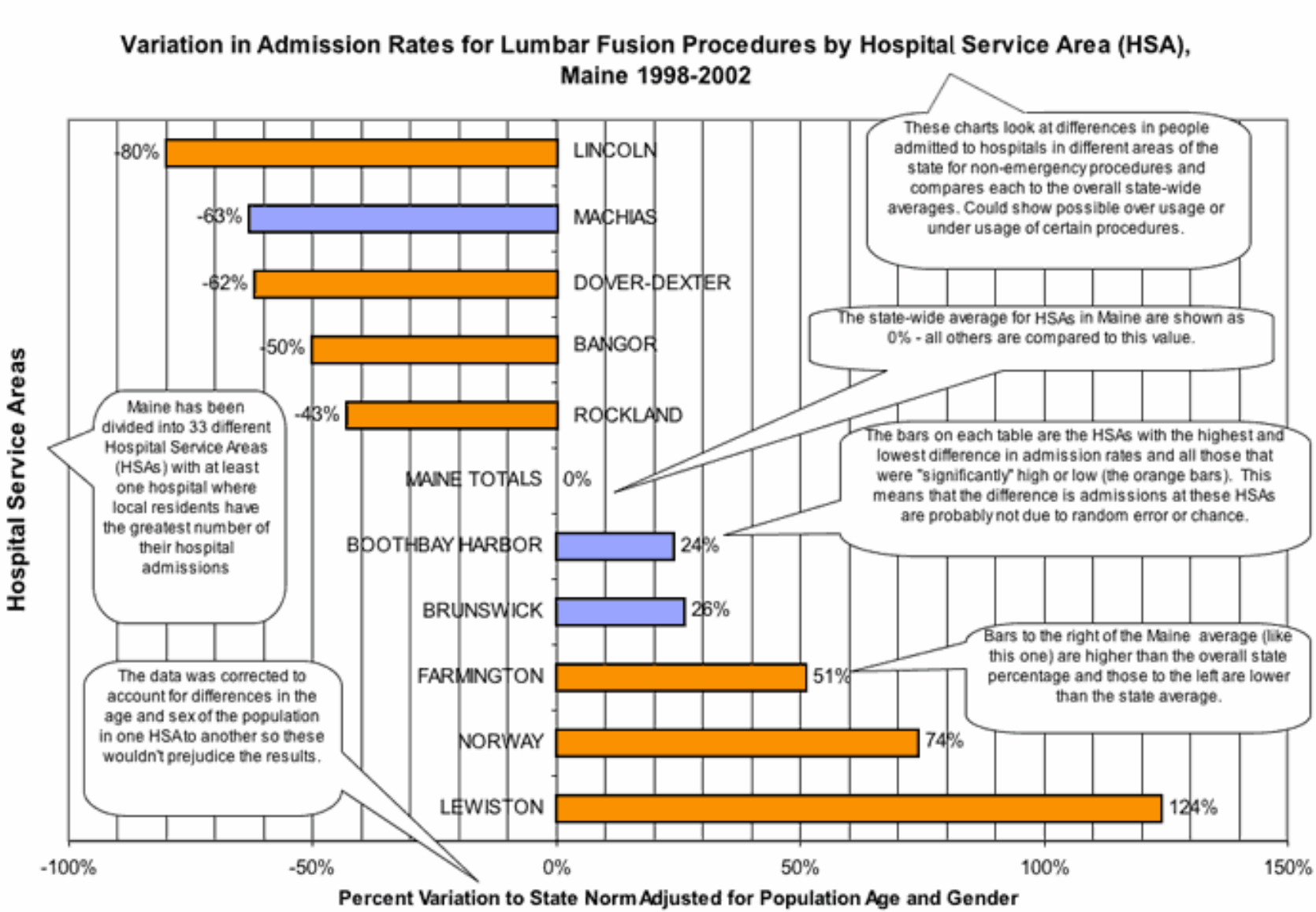
Source: NHS (UK), The Good Indicators Guide

Figure 4: An non-uniform system comparing NHS Trusts with each other

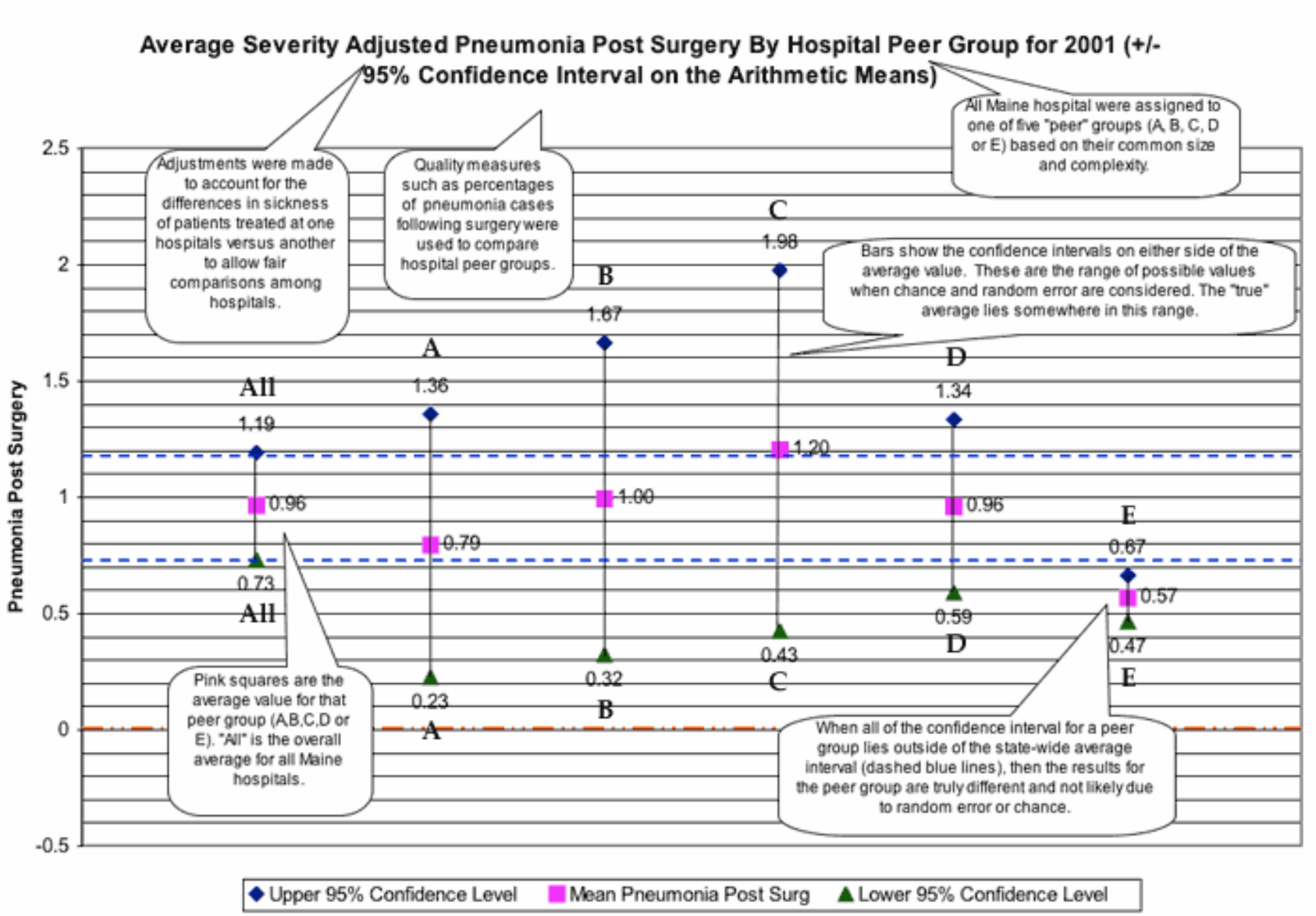


Source: NHS (UK), The Good Indicators Guide

Maine Quality Forum
Medical variation chart format



Maine Quality Forum
Quality analysis chart format



Indicator sets

- Indictor sets
  - limited yet comprehensive set
  - of coherent and significant indicators
  - that can be monitored over time and
  - disaggregated to relevant social units
- Agreed-on definitions
- Consider entire health field model
  - Engage a variety of stakeholders
  - Responsibility and accountability for performance
- Structure, process, and outcome measures
  - Evidence-based link between performance and health
- Established validity, reliability, and responsiveness to change
- Timely availability of data at a reasonable cost

Outline

- Principles of community health assessment and performance measurement
- Statistics of rates and proportions
- Balancing validity, reliability, and responsiveness to change
- ***Exercise***
- Privacy and confidentiality

Exercise

- Choose a DHHS District Health Profile
- Identify indicators that are especially strong and weak with respect to
  - Reliability
  - Validity
  - Responsiveness to change
- Diagnose the problems and suggest improvements

Outline

- Principles of community health assessment and performance measurement
- Statistics of rates and proportions
- Balancing validity, reliability, and responsiveness to change
- Exercise
- ***Privacy and confidentiality***

Confidentiality concerns due to small numbers

- Small numbers present concerns about the confidentiality of personal information
- Confidentiality protection
  - Ethical obligation of health officials
  - Maintain willingness of public and health care providers to participate in data-generating activities in the public interest

Confidentiality concerns due to small numbers

- Protection of individual medical and administrative records
  - Protected by state and federal regulations (HIPAA)
  - Information available to public health officials only for notifiable diseases and death
  - Public use micro data release protected by human subjects review boards
  - Generally not an issue for CHAs and performance measures

Confidentiality concerns due to small numbers

- Concerns for CHAs and performance measures
  - Discovering the identity of someone in a public health database
    - Can be harmful if it is a HIV registry, etc.
  - Discovering a person's personal or medical characteristics
    - Data tabulated by cause of death and income -- if identify of one case of lung cancer is known, so is his income
    - If one death in 50-59 year old men, system could disclose cause of death

Confidentiality concerns due to small numbers

- Specific measures currently in place vary substantially within and across states
  - Tend to follow rules set by data owners and same as for printed publications
  - State to state variation in legal protection of information on death certificates
- Some data are available in more detail to officials
  - Example: Minnesota student survey data
    - County data for available to the public
    - School-specific results available to school officials

Addressing confidentiality concerns due to small numbers

- Solutions
  - Suppress results based on less than *x* observations: “Rule of 5” or “Rule of 3”
    - Rule typically varies by data set
    - Suppress complementary categories
      - 10 AIDS deaths, 9 White ➔ 1 Black
    - Suppressing every row and column with a cell < x quickly leads to mostly blank tables
    - Can be defeated by recombination
      - Tabulate characteristics of people in Boston and Gosnold
      - Retabulate Boston alone, then subtract

Addressing confidentiality concerns due to small numbers

- Numerator rule
  - Combine “Other” with White or Black
  - Rule of 3 would prohibit publishing all but margins
  - Rule of 5 would not allow sex to be published

AIDS Deaths in County X				
	White	Black	Other	Total
Male	5/2000	4/200	1/2	10/2202
Female	2/2100	1/190	0/1	3/2291
Total	7/4100	5/390	1/3	13/4493

Addressing confidentiality concerns due to small numbers

- Denominator rule
  - Suppress results for populations less than *n*
    - E.g. *n* = 100,000
  - Based on assumption that there are a limited number of people with any given set of characteristics
  - Would prohibit publication of the following

Total deaths in a county of 90,000 population		
	White	Black
Male	350	150
Female	300	100

Addressing confidentiality concerns due to small numbers

- Numerator and event denominator rule
  - Suppress if a cell number subtracted from total with same characteristics < *x* (typically 10)

A: AIDS deaths, Black females aged 25-44	B: Total deaths, Black females aged 25-44	Release under numerator rule: A<6	Release under Missouri rule: B-A<10
1	100	No	Yes
6	100	Yes	Yes
1	7	No	No
6	7	Yes	No

Addressing confidentiality concerns due to small numbers

- Numerator and event denominator rule
  - Suppress if a cell number subtracted from total with same characteristics < *x* (typically 10)

A: Population	B: AIDS deaths, Black females aged 25-44	C: Total deaths, Black females aged 25-44	Release under denominator rule: A<100,000	Release under Missouri rule: C-B<10
95,000	1	15	No	Yes
95,000	1	2	No	No
105,000	1	15	Yes	Yes
105,000	1	2	Yes	No

Addressing confidentiality concerns due to small numbers

- Solutions
  - Restrict use to “trusted users” on a “need to know” basis
  - Top and bottom code/merge categories
  - Controlled statistical rounding and perturbation
  - Statistical modeling/smoothing
    - Data are replaced with a less variable estimate that doesn’t represent any one person’s data

Small numbers leading to confidentiality concerns

- GIS systems
  - At what point does a dot map disclose individual-level information?
    - How does this depend on the scale of the map?
  - Solution
    - Add noise – randomly shift point
    - Does this create a false individual record?
  - Geographical smoothing as disclosure avoidance
  - Needs more thought and analysis

Addressing confidentiality concerns due to small numbers

- Conclusions/recommendations
  - Data suppression rules
    - Not based in principle or theory
    - Tend to reduce utility of available information
  - Consider statistical modeling/smoothing
  - Federal and state public health agencies should reconsider goals and methods of confidentiality protection, guided by
    - Ethical and public health principles
    - Statistical theory

Big Ideas

- Indicator sets must
  - Address the appropriate range of community health issues and opportunities
  - Measure “well”: balance validity, reliability, and responsiveness to change
- “Small numbers” problem →
  - lack of statistical reliability
  - confidentiality concerns
  - unique and common solutions

References

- Evans RG, Stoddart GL. Producing health, consuming health care. In *Why are Some People Healthy and Others Not? The Determinants of Health of Populations*, Evans RG, Barer ML, Marmor TR, eds. Aldine De Gruyter, 1994.
- Institute of Medicine. *Improving health in the community: A role for performance monitoring*. Durch JS, Bailey LA, Stoto MA, eds. National Academy Press, 1997.
- Institute of Medicine. *The Future of the Public's Health in the 21st Century*. National Academy Press, 2003.
- Kindig DA. Purchasing population health: Paying for results. The University of Michigan Press, 1997.
- NCQA, *The State of Health Care Quality 2007*, 2008. Available at: http://ncqa.org/tabid/543/Default.aspx
- Stoto MA, Cosler LE. Evaluation. In Novick L, Mays G, eds., *Public Health Administration: Organization and Strategy for Population-based Management*, 2nd Edition, Jones and Bartlett, 2007.
- UK NHS Institute for Innovation and Improvement, *The Good Indicators Guide: Understanding How to Use and Choose Indicators*, 2008. Available at: www.apho.org.uk/resource/item.aspx?RID=44584

Data resources

- Maine
- DHHS District Health Profiles www.maine.gov/dhhs/boh/maine_dhhs_district_health_profiles.htm
- Maine CDC www.maine.gov/dhhs/boh/data_resources.htm
- Maine Quality Forum www.mainequalityforum.gov/mqsp01f.html

- Centers for Disease Control and Prevention
- National Center for Health Statistics www.cdc.gov/nchs/
- CDC Wonder wonder.cdc.gov/
- Behavioral Risk Factor Surveillance System www.cdc.gov/brfss/

- Other federal sources
- Agency for Healthcare Research and Policy www.ahrq.gov/data/
- Health Resources and Services Administration datawarehouse.hrsa.gov/
- Centers for Medicare & Medicaid Services www.hospitalcompare.hhs.gov
- National Cancer Institute www.cancer.gov/statistics/
- U.S. Census Bureau, American FactFinder
- factfinder.census.gov/home/saff/main.html?%20Lang=en

- Private foundations
- Annie E. Casey Foundation, Kids Count http://www.aecf.org/
- The Henry J. Kaiser Family Foundation, StateHealthFacts.org www.statehealthfacts.org/