Disease clusters: What's real, what's not, and how to tell the difference

Dick Clapp, D.Sc., MPH

July 7, 2009

Outline

- 1. What is a disease cluster?
  - Unofficial and official definitions
- 2. What are some typical ways evaluate reports of clusters?
  - CDC Protocol
- 3. What are some statistical methods?
  - Woburn and Cape Cod studies
- 4. How do you tell what's "real"

Location and Disease

- location = clue to cause
- history (e.g., Snow's Cholera studies)
  - Migrant studies (e.g., breast cancer in Japanese women who move to the US)
- contemporary examples
    - Lyme disease, infectious diseases (H1N1 '09)
    - heart disease and "Mediterranean diet"
    - cancer in Woburn, Tom's River NJ,
    - Birth defects, scleroderma, polycythemia rubra vera

Terms of the debate

AMERICAN JOURNAL OF EPIDEMIOLOGY

Vol. 132, Suppl., No. 1
Printed in U.S.A.

Keynote Presentation

A SOBERING START FOR THE CLUSTER BUSTERS' CONFERENCE

KENNETH J. ROTHMAN

Epidemiology
1995

Should We Boost or Bust Cluster Investigations?

Daniel Wartenberg

MEDICAL DISPATCH

THE CANCER-CLUSTER MYTH

When a dozen people in a neighborhood develop tumors, it can't be coincidence. Or can it?

BY ATUL GAWANDE

Some of the issues

- typical kinds of cluster studies
  - "too much cancer in my neighborhood," "lots of kids with autism near the plastics factory"
- epidemiologic aspects:
  - small-scale, low statistical power
  - non-communicable diseases (usually)
  - long latency (sometimes, e.g., cancer)
- suspected environmental cause

Cluster Epidemiology

: the statistical analysis of spatial or space-time distributions of disease, with the goal of "detecting clusters."

- local data compared to standard population
- p-values; significance testing
- significant clustering means a cluster is detected, or follow-up necessary
- "pre-epidemiology"

“What is a cluster?”

- CDC (2008):
  - “a cluster of public health concern is defined by an unusual aggregation, real or perceived, of health events in time and space that is reported to a health agency”
- Dictionary:
  - “a group of the same or similar elements occurring closely together; a bunch (like grapes)” [American Heritage Dictionary]
- Street:
  - “if it looks like a duck, walks like a duck, and quacks like a duck, it’s a duck.”

Cluster Modifiers

apparent perceived

putative actual

true suspected

possible reported

circumstantial

significant

Cluster

Source: Jacquez et al.

Two basic concepts of “cluster”

- Cluster_1 (cluster-as-excess)

  : an observable excess of cases in space-time.

- Cluster_2 (cluster-as-causal-excess)

  : an observable excess of cases in space-time due to a cause of interest.

Rothman
"Sobering Start…"

- Cluster studies:
  - small numbers
  - inhomogeneous case definition
  - biased ascertainment
  - latency
  - poor information on exposures, etc.
- Conclusion

  "no new statistical methodologies are needed to refine our study of disease clusters…"

Cluster Epidemiology vs. Spatial Epidemiology

- Cluster Epidemiology
  - statistical analysis
  - p-values
  - significance testing
  - seeks to detect “clusters”
  - issue of boundaries, etc.
  - faith in statistical testing

- “Normal” Epidemiology
  - data description
  - measuring disease incidence vs. location
  - identifying related causes
  - analogy: occupational epidemiology

CDC Protocol (MMWR, July 26, 1990/39(RR-11);1-16)

- 1. Initial contact and response
  - Gather information from concerned citizen
- 2. Assessment
  - Preliminary calculations, case verification, literature review, feedback to advisory committee
- 3. Major feasibility study
  - Assess potential study design, logistics, cost

CDC Protocol (continued)

- 4. Etiologic Investigation
  - Develop protocol and conduct study
  - Expected to “contribute to epidemiologic and public health knowledge”
- Examples include Woburn childhood leukemia follow-up, Tom’s River childhood cancer study, Camp Lejeune studies, many others over past 20 years

New England example

- "A suspected cancer cluster is more likely to be a true cluster, rather than a coincidence, if it involves:
  - A large number of cases of a similar type . . .
  - A rare type of cancer . . .
  - An increased number of cases of a certain type in an age group that is not usually affected by that type of cancer."

New England example (cont.)

- “statistical significance will determine if additional steps are required . . . The concerned individual or community is provided with an official letter. . .
- “Because a variety of factors often work together to create the appearance of a cluster where nothing abnormal is occurring, most reports of suspected cancer clusters are not shown to be true clusters.”

Case Examples

- Childhood leukemia and contaminated drinking water in Woburn, MA
- Breast cancer on Cape Cod and in Massachusetts

Woburn Citizens' Concerns

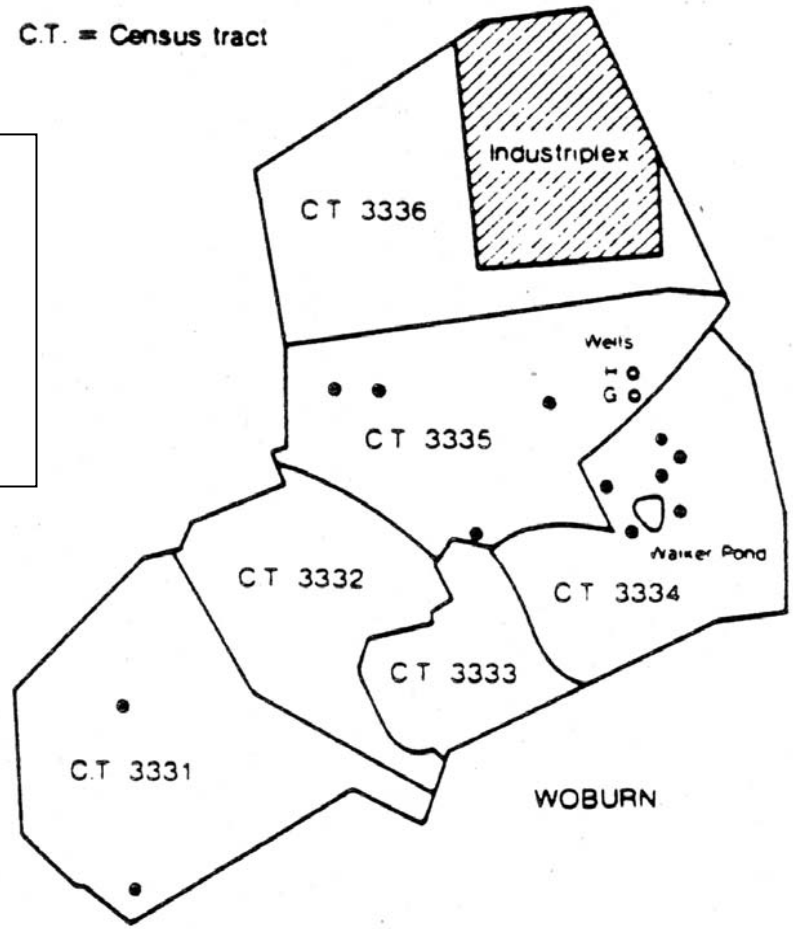
- Years of toxic waste dumping from leather tanneries, chemical plants
- Strong odors in East Woburn
- Water tasted bad periodically

- Apparent cluster of childhood leukemia in East Woburn neighborhood
- Local and State officials unresponsive

Citizen Action

- Created grassroots organization called FACE (For A Cleaner Environment)
- Called for health and environmental studies
- Demanded clean-up of contaminated properties
- Several families filed lawsuit (subject of book and movie, “A Civil Action”)

Woburn Childhood Leukemia Cases 1969-1979



Results of 1981 Study

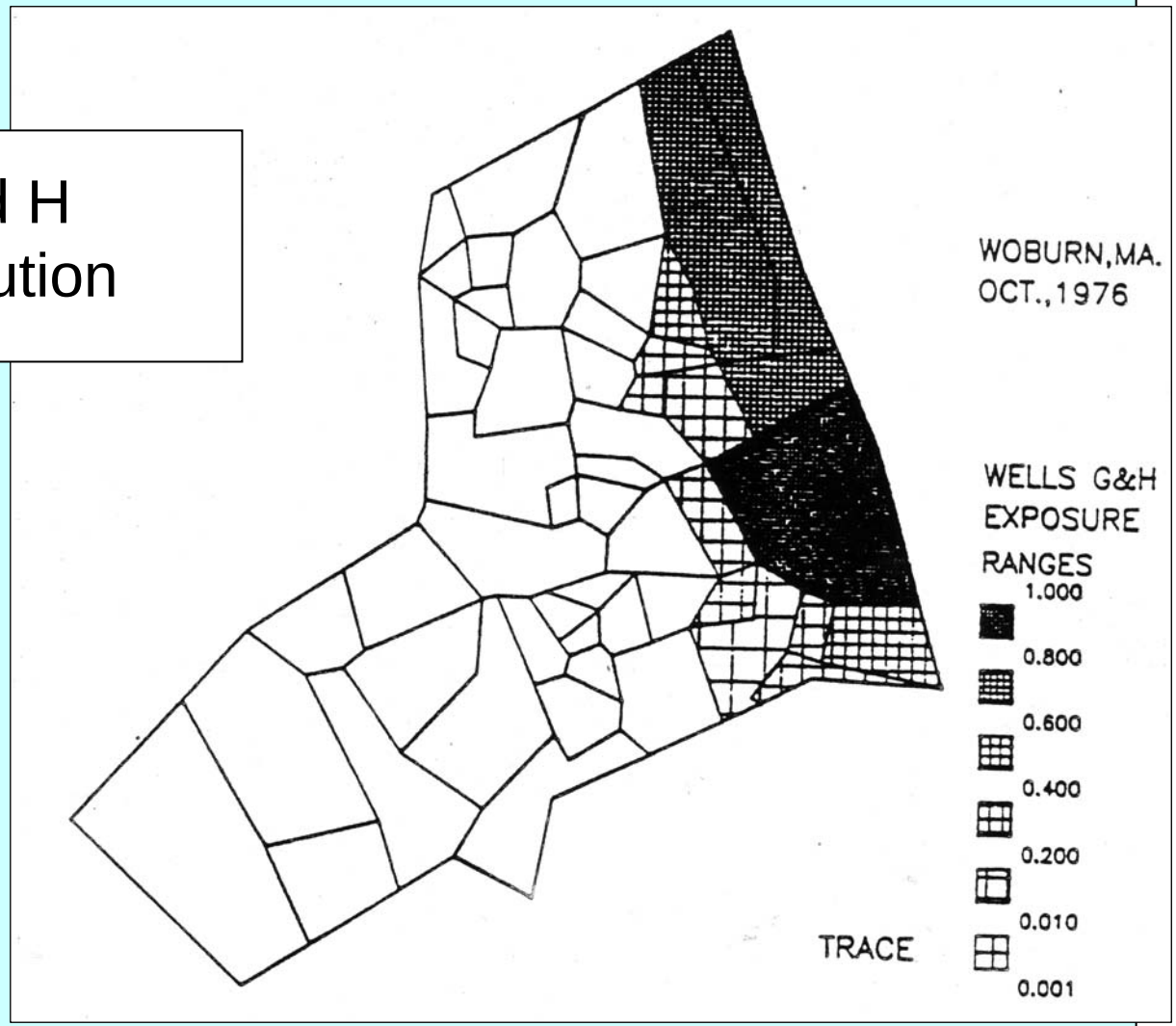
- Case-control study of 12 childhood leukemia cases diagnosed 1969-1979
- No association with any particular exposure, family history, medical history
- Elevated Standardized Incidence Ratio SIR=2.3, p=.007 (see Cutler JJ, et al. Childhood Leukemia in Woburn, Massachusetts. *Public Health Reports* 101(2):201-205, 1986)

Harvard/FACE Study

- Citizens and researchers conducted town-wide telephone survey
- Used Woburn water distribution model from Dept. of Env. Quality Engineering
- Estimated exposure to contaminated wells G & H was two-fold higher in childhood leukemia cases diagnosed 1964-1983

(see Lagakos SW, et al. *J Am Stat Assoc* 81:583-596, 1986)

Wells G and H Water Distribution

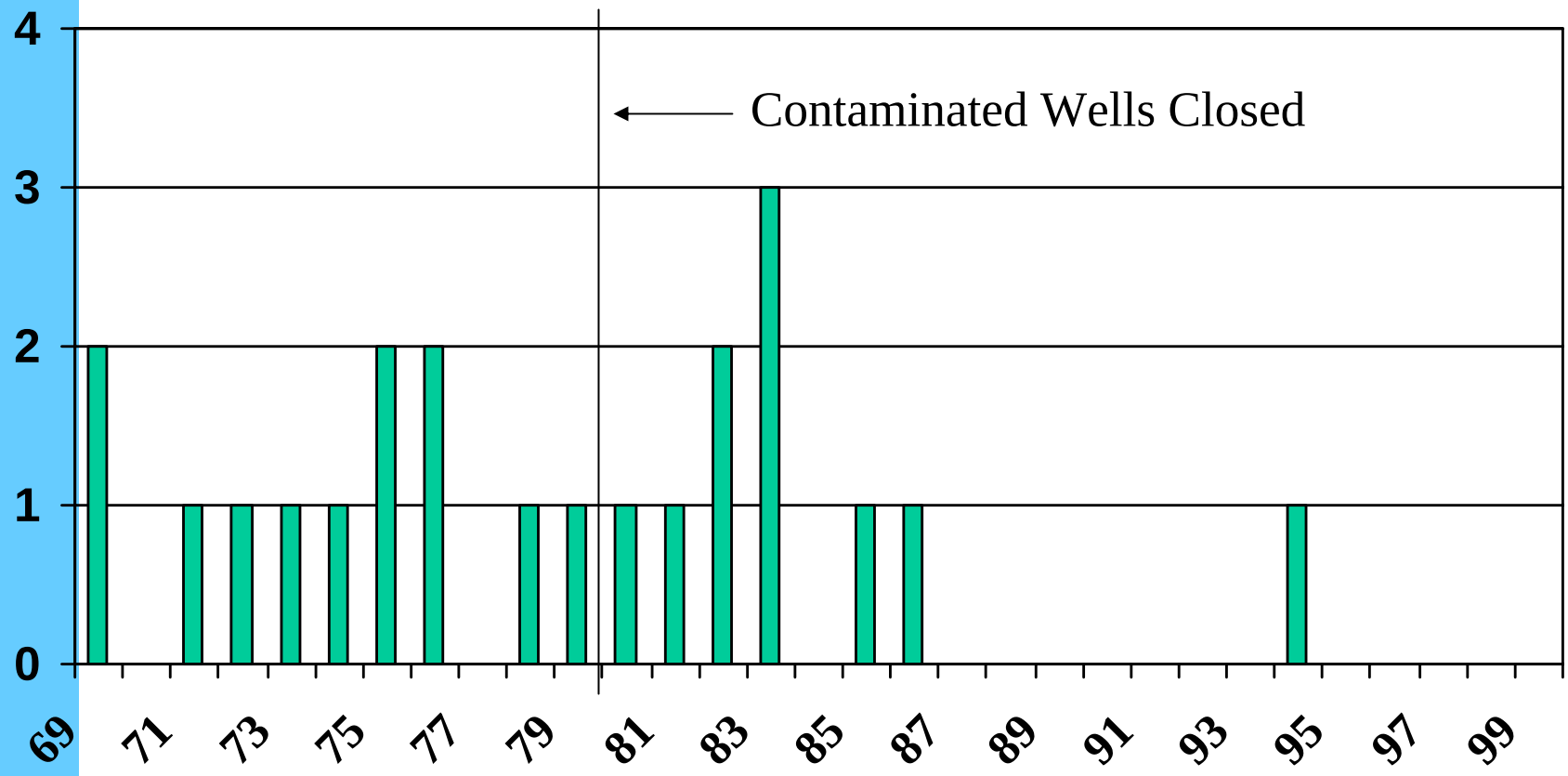


DPH Follow-Up Study

- Case-control study of childhood leukemia diagnosed 1969-1986 (Costas, et al., 2002)

- Used detailed water distribution model with estimates of G & H exposure by month
- Increased risk of leukemia in children whose mothers exposed during pregnancy (OR=8.3; 95% C.I.=0.7-94.7, p(trend)<.05)

Number of Cases of Childhood Leukemia Diagnosed by Year of Diagnosis in Woburn, MA Jan. 1, 1969 - Aug. 1, 1999



Groundwater Treatment Process

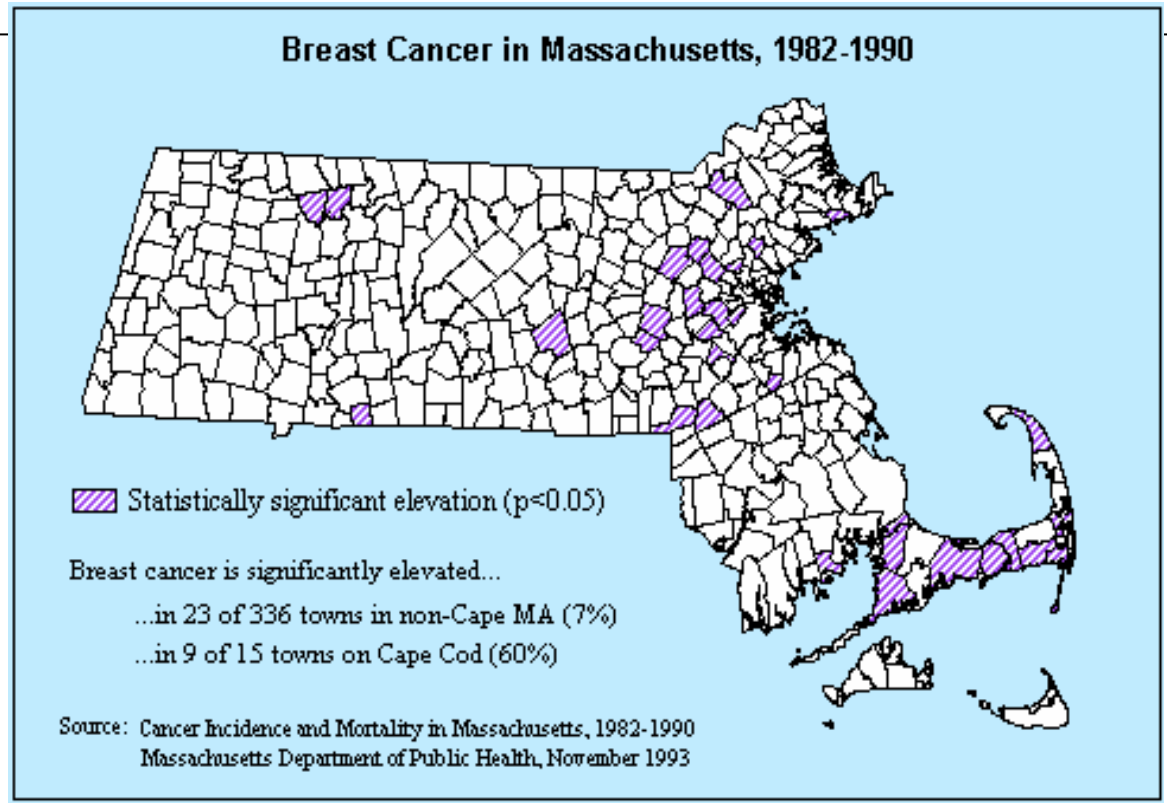
Statistical methods

- Standardized incidence ratios for population of interest
  - Compare observed to expected cases, calculate p-value and 95% confidence interval
- Utilize cluster statistics
  - CDC software “Cluster version 3.1”
  - Knox space-time statistic
  - Scan statistic, cusum statistic, etc.
- Kulldorff methods “SaTScan v8.0.1, June, 2009”

Breast cancer in Mass.

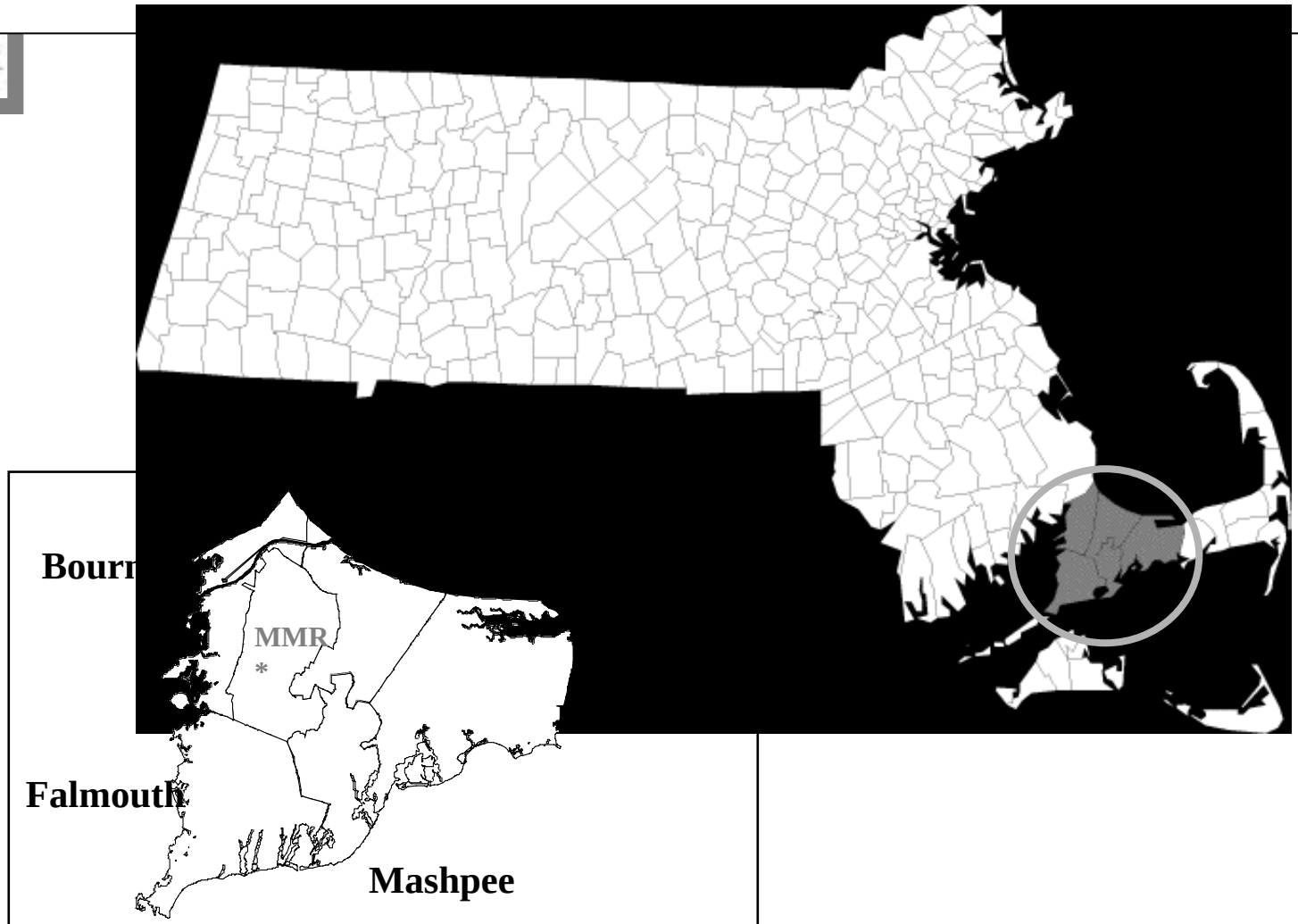
- Silent Spring Institute/BUSPH studies
  - Mass. Breast Cancer Coalition's critical role
  - Initial descriptive epidemiology
    - SIR significantly elevated from 1982 to 1990
  - Upper Cape study
    - Breast cancer in relation to PCE water contamination, Mass. Military Reservation exposures

Breast Cancer SIRs



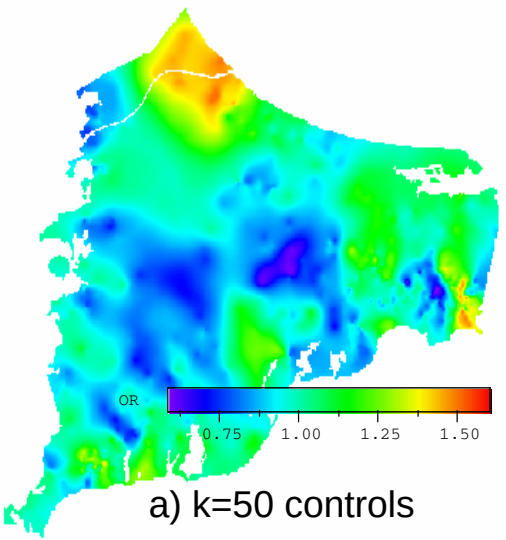
- source: Silent Spring Institute, www.silentspring.org

Upper Cape Cod

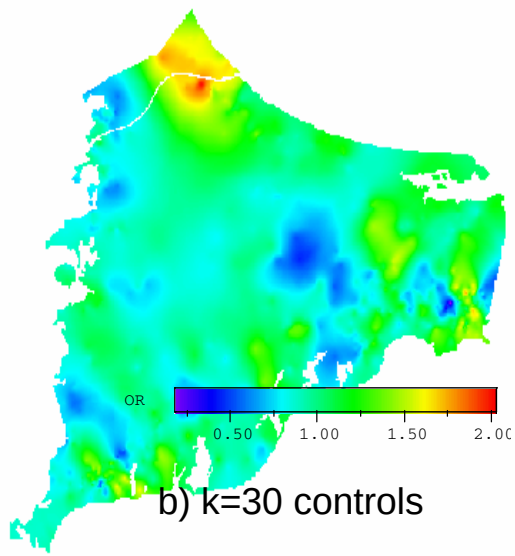


* Massachusetts Military Reservation

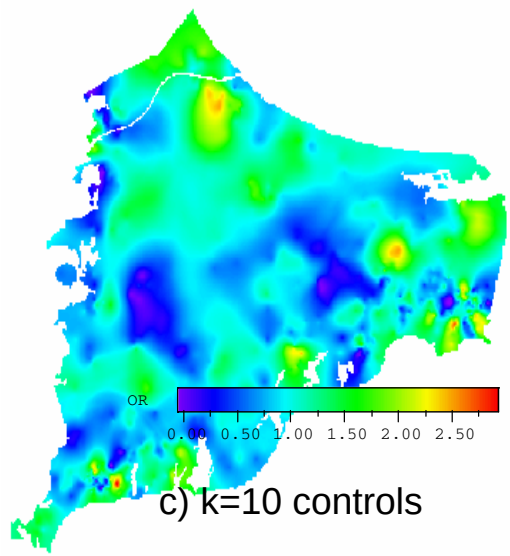
Adjusted breast cancer ORs - Smoothed (Paulu, et al., 2002)



a) k=50 controls

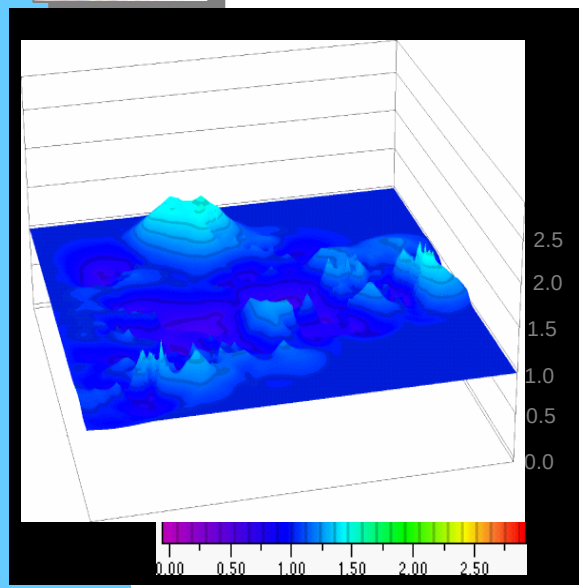


b) k=30 controls

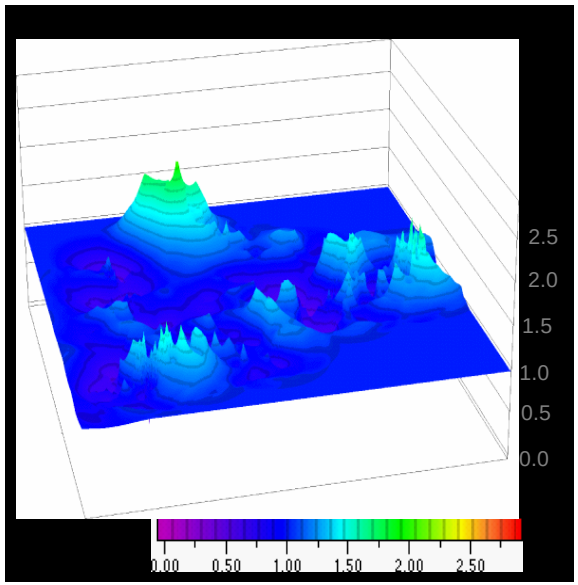


c) k=10 controls

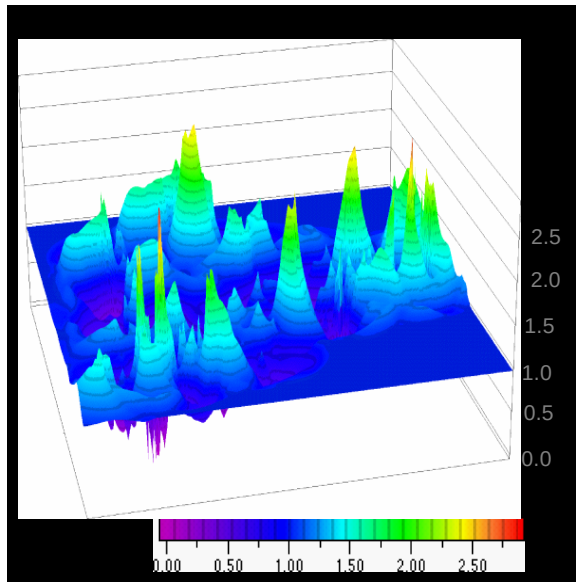
Adjusted ORs in 3-D



a) k=50 controls

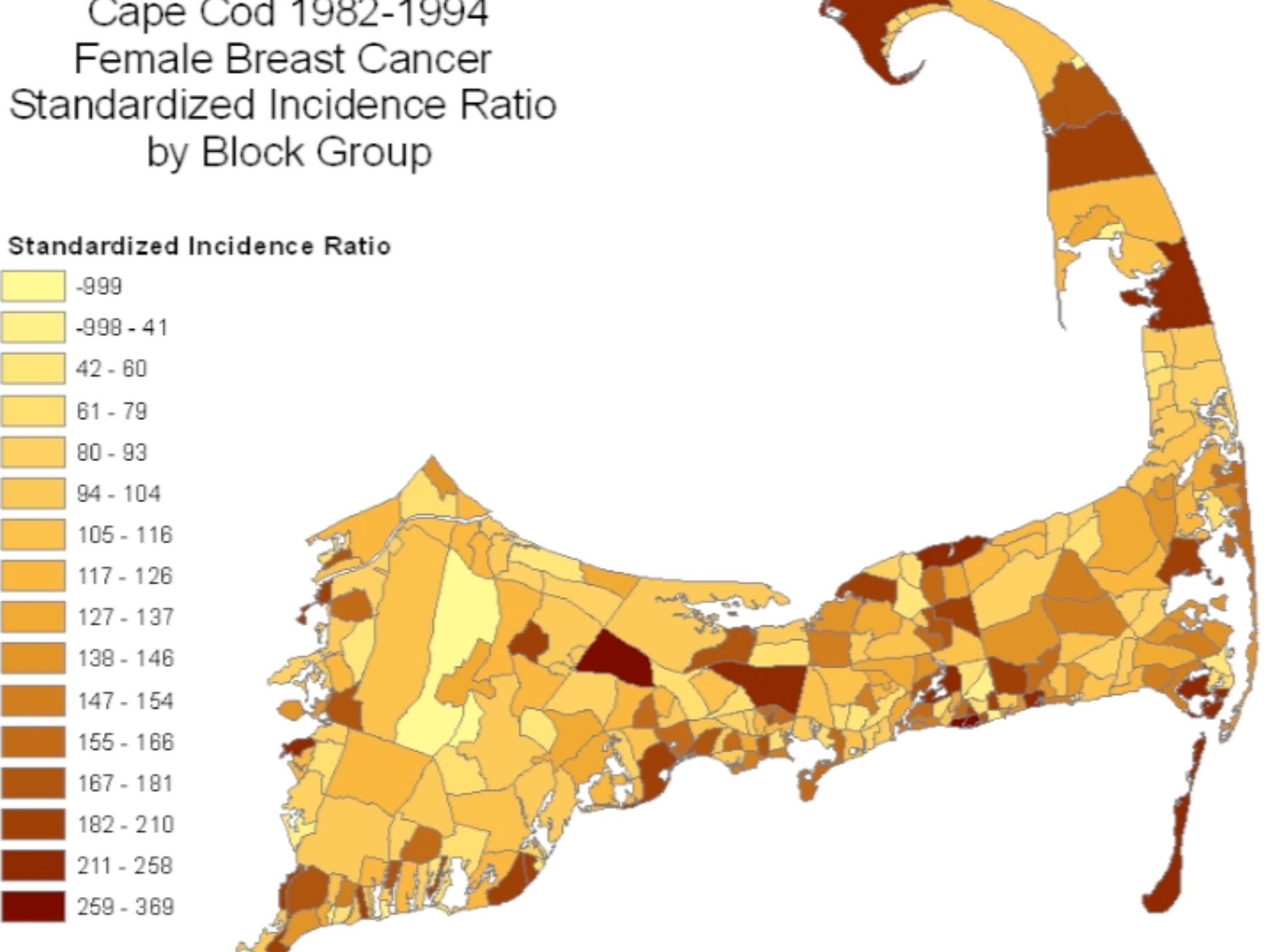


b) k=30 controls



c) k=10 controls

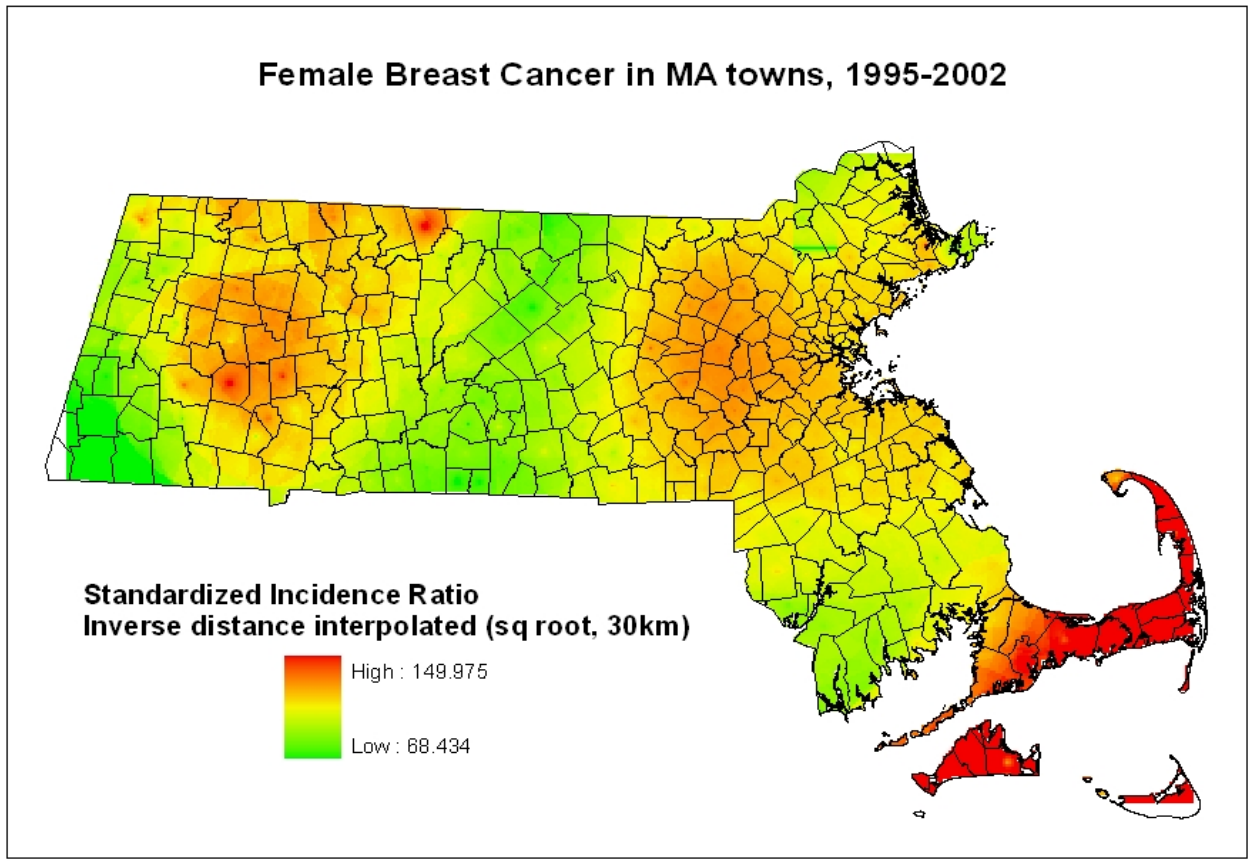
Cape Cod 1982-1994 Female Breast Cancer Standardized Incidence Ratio by Block Group



Breast cancer in Mass. (cont.)

- On-going studies of household, consumer product and environmental exposures
  - Silent Spring Institute results at www.silentspring.org
  - U. Mass.- Lowell/TURI/Silent Spring collaboration at http://gis.uml.edu/mediawiki/index.php/Breast_Cancer_Risk_Project
  - BUSPH expanded case-control and cohort studies, further GIS methods (Vieira, et al.)

Interpolation of town-level breast cancer incidence data (Brown, et al., 2009)



Source: SIlent Spring Institute

Inverse root distance weighting was used on town-level aggregated SIR data on town geographic centroids, with a radius of influence of 30km. (Data: SSI; map: TURI)

How to tell the difference?

- No “magic wand” can tell whether a cluster has a real cause or not
- Basis for reasoned judgment
  - Toxicologic and epidemiologic literature
  - Assessment of local disease and exposure patterns
  - Interpretation of detailed epidemiologic studies, if available

The obligation to act

- “From the right to know and the duty to inquire flows the obligation to act.” (Sandra Steingraber, *Living Downstream*, 1997)
- Cluster investigations do not take precedence over reduction of likely harmful exposures; no proof of harm is not proof of no harm

Exercise

- Sandwich, MA is adjacent to the Massachusetts Military Reservation (MMR) and has been part of an on-going series of investigations of cancer and adverse birth outcomes over the past two decades. Recently, parents have raised a concern about a possible cluster of Ewing's Sarcoma (a rare bone cancer) in Sandwich children. They have documented two cases in Sandwich and five additional cases in other Cape Cod communities.

Exercise, continued

- As an employee of the State Health Department, what would you recommend as a way of evaluating the community's concerns? What types of data would you look at in order to get an initial assessment of the problem? Consider the exposures that have already been identified in previous studies (see the MMR website, publications by Aschengrau, et al., and other sources you may find). Look at the Mass. Cancer Registry on-line data and consider what additional data you would want to request.