TOP TEN LIST

TEN BEST WAYS TO DO SPATIAL ANALYSIS OF HEALTH DATA BADLY

With apologies to David Letterman, and thanks for editorial assistance to Elizabeth Kirby and for their insights the following colleagues:

Russell Rickard, CO Dept. of Health and Environment
Sara McLafferty, Univ. of Illinois
Geoffrey Jacquez, Biomedware, Inc.
Sam Lefevre, UT Dept. of Health
Tom Talbot, NY Dept. of Health
Jason Salemi, Univ. of South Florida
Gonzalo Vazquez-Prokopec, Emory Univ.

R.S. Kirby, February 2012

Number 10

Scales Are Only Used for Weighing Things

Ignore spatial scale – it doesn't matter anyway.

In fact, administrative units don't matter either, so use the most convenient source, such as self-reported county of residence. Knowing the county in which a person resides is just as informative as knowing their street level data. In a pinch, use the city name associated with the post office address.

Doing both of these will almost guarantee flawed and invalid results.

Number 9

It Was On Fire When I Stepped On It!

Proximity IS exposure. The closer a case is to a point source, the more exposure is occurring, now and in the past, regardless of what environmental modeling or fate might suggest.

Never worry about the confounding effect of environmental justice, especially when studying socially stigmatized health outcomes.

Number 8

Brother, Can You Spare a Dime?

Always remember, all rate transformation methods that include the word ‘Bayesian’ are suitable to your data.

Borrowing strength (empirically) from neighbors does not necessarily mean that disease rates will be stronger.

Number 7

Clusters Anyone? Also Available as Jewelry, Peanut Candy, Breakfast Cereal . . .

All spatial clusters are perfectly circular in shape, centered directly on the putative cause.

Never analyze your data for variation in ascertainment or data quality – areas served by hospitals with excellent reporting of cases are clusters because they have higher rates.

Number 6

No One Has Ever Studied This Before

Never adjust your rates for covariates. How could variations in age, race/ethnicity, sex possibly affect spatial patterns of incidence or mortality?

Better yet, never test for risk factors individually, Control for everything. Use the maximum number of strata for each variable, especially continuous variables. The least parsimonious model should be the correct one as it considers everything.

Number 5

Do You Mean to Say This is Scientific?

All analyses should be conducted using the relative concept of space. Conduct all analyses using Euclidean distances (e.g. buffers, kernel densities, centroids), on unprojected data.

You'll be sure to get interesting results, and if they are published, you can explain what they mean at your next job interview.

Number 4

Four Blades are Better Than Three

Smooth your data, then analyze for spatial autocorrelation. What a surprise, you are sure to find it! So now, over-interpret it and confute causal factors with the spatial autocorrelation introduced by the smoothing algorithm.

Always assume that spatial stationarity exists. Just include a disclaimer at the end of your report that you didn't test for it, and if doesn't exist your results might be invalid if the assumptions are too far off from the reality of your data.

Number 3

Is There a Story Behind the Numbers?

Use spatial techniques that ignore numerators and denominators. Who cares about stability of rates anyway?

A few helpful hints:

1. Never present maps of uncertainty along with your research findings.
2. Small numerators that lead to large rates should be the focus of any report or presentation.
3. Data suppression is for fools.
4. Data suppression can make you look like a bureaucratic genius.

Number 2

Where No One Walks, Alone

Spatial health data analysis is best done by you alone.

Don’t collaborate with spatial statisticians, medical geographers, or epidemiologists. And most definitely, don’t include local content experts on your research team. Treat this work as an individual sport, and you’ll be almost certain to do spatial analysis of health data badly.

Ten Best Ways to Do Spatial Analysis of Health Data Badly

Number 1

So That's What Manuals are For!

For the best results, always use complex methods straight out of the software, with the default settings for all parameters and selections.

Who cares about the details – if it's good enough for the software engineers it should be good enough for you.