Field Guide of Contextual and Relational Epidemiology

Jon Zelner

2023-09-08

Table of contents

1 Inspirations and Influences

2 Inspirations and Influences

What - if any - kind of book needs to be written in the realm of contextual epidemiology? And why would I decide to write something like this?

2.1 What is the problem to be solved?

There seems to be a hole in the literature on spatial epidemiology that comes to grips with the human aspects of risk. This is a particularly hard problem in infectious disease epidemiology, where the role and importance of context is increasingly understood, but it has an uneasy coexistence with the mainstream of the field.

In many ways, my goal in writing this is selfish, to satisfy a personal curiosity I suspect is of some greater import to at least a few people. Specifically: What does it mean to do *contextual* epidemiology? Is this a meaningless term that just appeals to the social and natural science centers of my brain, or is it tapping into something more meaningful?

Part of my inspiration is a reaction to the technocratic impulse and imperative in modern epidemiology. In particular, I find myself a bit paralyzed by the worry about what happens when we operate under the assumption that escalating methodological complexity is an imperative and that the road out of socio-epidemiological problems is paved with technological solutions.

On the other side, simply put, I love the methodological tools of spatial epidemiology, Bayesian hierarchical analysis, and systems modeling. I have learned more than I ever could have hoped through these learning, tinkering with, and applying them to problems both in the real world and in my own head. But for me, the large majority of lessons learned have been from their conceptual isomorphisms (or conflicts) with the world as it appears to us through qualitative and quantitative data, rather than in their parameter values and predictions.

For me, this book is about resolving this cognitive dissonance while providing useful 'how-to' pointers along the way. I hope to articulate the affirmative case for a systems-based, contextually-sensitive, justice-oriented, morally and ethically opinionated, and theoretically driven approach to epidemiology. Along the way, I hope to show why the tools of such an approach are necessarily heterogeneous in nature and require us to accept uncertaintities quantitative and epistemic.

2.2 What are the elements of contextual epidemiology?

I arrived at the idea of contextual epidemiology from spatial and hierarchical analysis. The stereotypical challenge in this setting is to adjust away the impact of context to get at some more generally meaningful parameter estimate, e.g. a treatment effect. But another way of thinking about this is that the hard work is in characterizing not only the 'main' or fixed effects, but in capturing the drivers of variability in outcomes across locations.

But not every contextual question is strictly spatial: If we care about how the structure of social networks impacts individual and collective risks, we are talking about context once again. In the network example, the context is one's network 'neighborhood', the collection of individuals one interacts with. In reasonably homogeneous networks, groups of individuals may share very similar or almost identitical network contexts.

Ultimately, the key element of context is *relatedness*: These relationships may be micro-level social relationships or macro-scale spatial ones, but they may also be temporal in nature. Temporal relationships may occur at a micro scale, e.g. the rise or fall in incidence of an infectious disease during a given week is likely to be a function of the prevalence of that same disease in the previous week. But these temporal relationships are also often more macro-scale and historical in nature: The long history of racial residential discrimination in the United States is undoubtedly a driver of many racial health dispartities we see today.

Making sense of the types of patterns we see in the real world requires us to first identify:

1. A question we want or need to answer.
2. The most important types of relationships impacting our outcome of interest (time, space, individual-to-individual).
3. A methodological approach that will allow us to characterize the impact of those relationships on the outcome we care about.

It is important to note that the choice of method comes *last* here: A key motivation of this book is to sidestep the tendency to train ourselves into methodological hammers looking for data nails to whack away at.

In addition to this, there is no pre-supposition that the appropriate method is necessarily 'fancy' in the sense of being conceptually or mathematically complex, computationally intensive, or even primarily or at all quantitative in nature. Whether this is the case is entirely a function of what happens at the intersection between question, data, and theory.

2.3 A book, not a cookbook

> "A common mistake that people make when trying to design something completely foolproof is to underestimate the ingenuity of complete fools." - Douglas Adams, Hitchhikers Guide to the Galaxy

2.4 Influences

There are any number of books and papers out there that have articulated a similar perspective on the tools of quantitative analysis. The following have been particularly important for my own thinking and their influences will be felt throughout this work:

- Statistical Rethinking (1)
- The Ecological Detective (2)
- ARM/Regression and Other Stories (3)

What makes these works so useful, strong, and enduring is the way that they articulate a coherent, opinionated perspective on the meaning and use of a set of methodological tools. On top of that, they are engaging and fun to read - the sort of thing you return to over time not just to get specific methodological tools, but to be exposed to their perspective.

References

Part I

Introduction

3 An invitation to relaitonal epidemiology

One theme that comes up consistently when I'm teaching and talking to students and others about spatial or relational epidemiology is that it is an aura of inaccessibility. This makes a certain kind of sense: Making maps and measuring and modeling spatial relationships might seem like it is outside of the classic analytic toolkit of working epidemiologists, not to mention other fields in public health, medicine, and the social sciences. In fact, this was my take on it as well, until I realized space has always been at the heart of my research, even when I didn't realize it!

The diagram in Figure 3.1 reflects my interpretation of the way someone coming to this area of research and practice for the first time might see it:

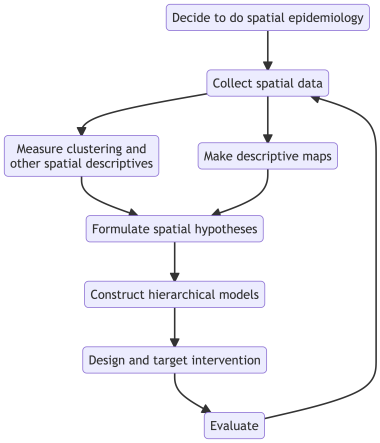

Figure 3.1: A textbook version of the spatial epidemiology workflow, without all the twists and turns that characterize actual research.

In this view, spatial epidemiology or analysis is taken to be a set of relatively fixed and well-described ideas and procedures connected to a highly technical set of methods. When we come from this perspective, starting a spatial project requires us to embark with a backpack that is already filled with specialized spatial tools.

3.1 You are (almost certainly) a spatial analyst already

My goal in this short essay is to chip away a bit at the idea of spatial epidemiology as something separate and apart from mainstream epidemiology and health. Instead, I argue that it is better understood as an emergent phenomenon that acts as a loose wrapper around a core set of ideas and tools that are part of the working arsenal of most professionals, students, and researchers in public health.

To explain what I mean, and why I think it's important, I'm going to subject you to a bit of autobiography about my own intellectual and professional trajectory.

3.2 Maps: Gateway or destination?

I began working as an epidemiologist or epidemology-adjacent type as a PhD student at the University of Michigan some time around 2006. As part of my dissertation research, I worked with Joe Eisenberg on an analysis ofthe role of social networks as sources of risk and protection against diarrheal disease and other infections in a group of villages in an area of rural Ecuador:

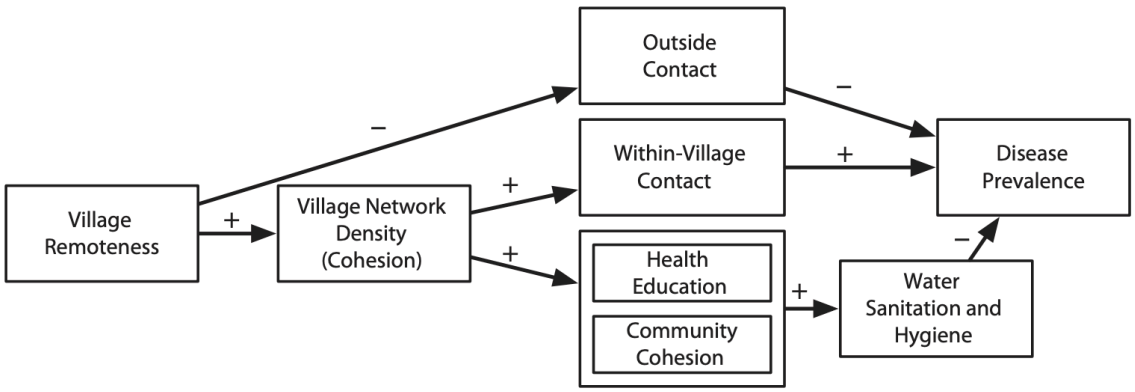

Figure 3.2: A figure from my dissertation research representing hypothesized relationships between village context (represented by inaccessibility or 'remoteness') and variation in disease outcomes within and between villages. (From (**zelner2012?**))

Looking back, this was indisputably a spatial analysis: We were interested in how local social contexts impacted variation in health outcomes *across* a set of 20+ villages and also how within-village variation in social connectivity impacted risk *within* villages.

We employed multi-level data about common characteristics of individual villages, households, and the individuals within them. But at this time, I thought of myself as doing a few things, but none of them were spatial:

1. Infectious disease epidemiology: Why and how do people become *infected* with various pathogens?
2. Social epidemiology: How do social *relationships* impact disease outcomes?
3. Network analysis: How does the *structure* of relationships impact individual and community health?

As it happened, all of these things were correct. But what I didn't really understand at the time was that the collection of these different approaches into a single analysis made it spatial or geographic in nature, even if I didn't realize it

3.3 A process of progressive revelation

As a postdoc, working with Ted Cohen, I began analyzing data from a large study of household-level tuberculosis transmission in Lima, Peru. Figure 3.3 illustrates the model we developed to characterize household-level differences in transmission rates as a function of exposure type:

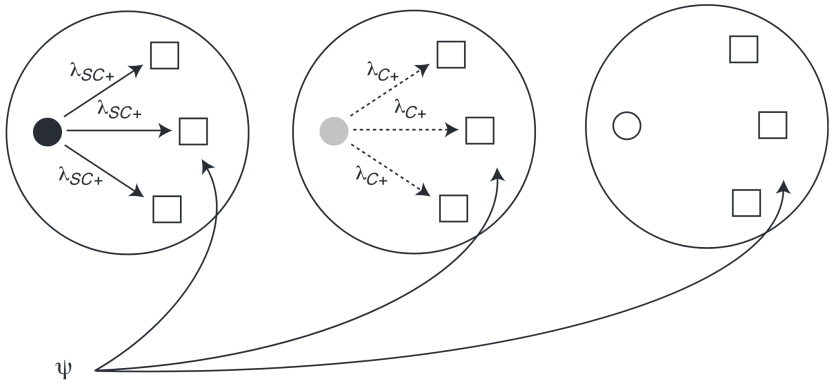

Figure 1. The figure illustrates the different sources of tuberculosis infection in the infection risk model. Smear-positive/culture-positive index cases (black circle) are hypothesized to be the most infectious, followed by smear-negative/culture-positive index cases (gray circle) and then smear-negative/culture-negative index cases (white circle). λ_{SC+} and λ_{C+} indicate the risk of infection for an uninfected household contact (white square) exposed to a smear-positive/culture-positive index case or smear-negative/culture-positive index case, respectively. ψ is the community risk of infection to which all households are subject. The solid arrows indicate a higher hypothesized level of infectiousness than values represented by the dashed arrows. Random intercepts are included in the model at the household and health center level to account for correlated responses within these units.

Figure 3.3: Characterizing household-level variation in risks of infection from community vs. household exposures (Figure from (**zelner2014?**))

At the time, I knew that these households were distributed across neighborhoods of Lima, but I didn't give it much thought. I was more interested in risks experienced by an average

household. And to be honest, I didn't know that spatial metadata were available on each of the households, since I wasn't involved with the data collection!

In the interim, I got the chance to work on some collaborative projects with an explicitly spatial focus. In one, we reconstructed an outbreak of morbillivirus (think: measles) among a herd of migratory dolphins (Figure 3.4).

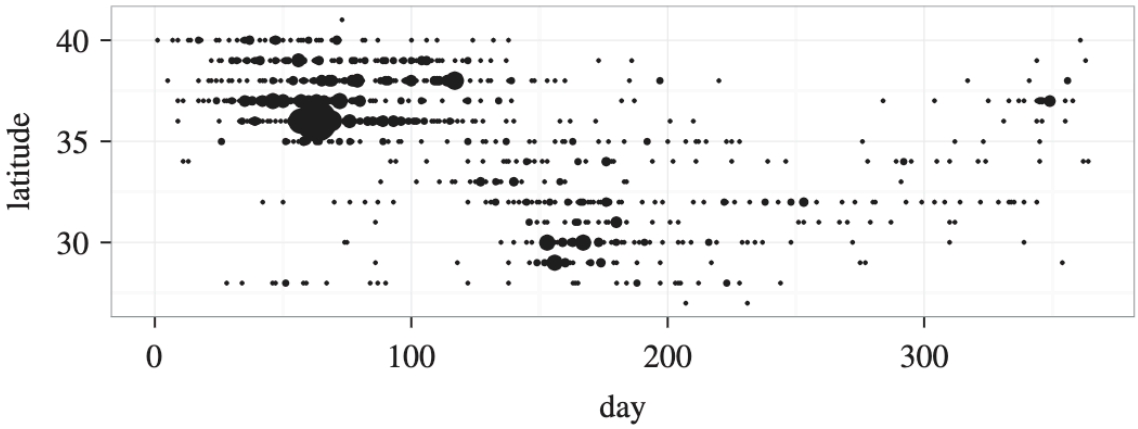

Figure 3.4: Locations of dolphin strandings during a morbillivirus outbreak in the North Atlantic (dot size indicates a greater number of strandings; Figure from (4))

In another, we looked at the relationship between environmental risks, such as neighborhood-level flooding, on the rate of pediatric diarrheal disease in Ho Chi Minh City in Vietnam (Figure 3.5).

These were the first experiences I had explicitly looking at these outcomes as a function of geographic space. While I had previously thought that mapping and spatial analysis and health geography were big scary things I couldn't do, I started to realize something important: These projects were not substantively very distinct from ones I had done before. The difference was that we were explicitly talking about spatial relationships and making maps (or simple one-dimensional diagrams as in Figure 3.4), instead of implicitly as in Figure 3.2 or Figure 3.3.

After completing these other projects, I dove back in to the Lima TB data to look at the drivers of multi-drug resistant TB (MDR-TB) risk. This was when I finally found out (some 2 years after I had started working with these data!) that spatial information on each household was available. So, with great trepidation, for the first time I made a map to explore spatial variability in MDR-TB outcomes.

And when I did this, we instantly saw that there were seemingly meaningful differences in the rate of TB overall, and MDR-TB in particular, across different health center catchment areas:

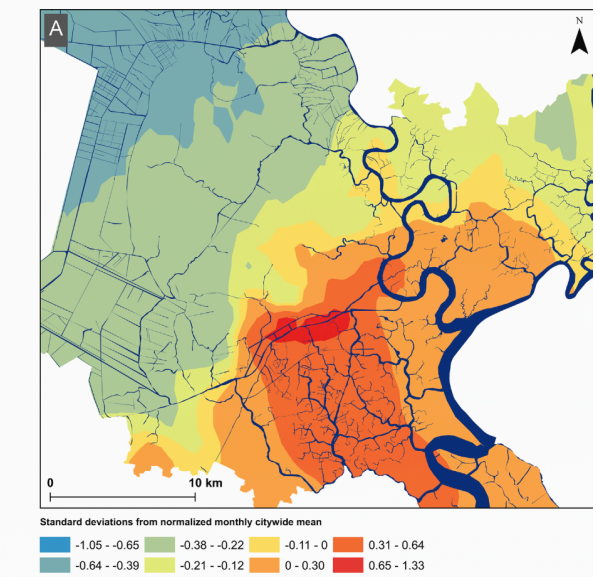

Figure 3.5: Incidence of pediatric diarrhea across neighborhoods of Ho Chi Minh City, Vietnam (Figure from (5))

This was the moment, some 10 years after I dipped my toes into the world of infectious disease epidemiology, where I realized I had been doing spatial work all along.

3.4 So what?

Why am I bothering you with this tedious and indulgent bit of personal history? *It's because it took me way too long to recognize that spatial epidemiology was a wrapper around a set of skills and ideas I had been working with for many years before I recognized what I was doing.* I was intimidated by anything preceded by 'spatial-': it sounded like a bunch of skills I didn't have and couldn't acquire.

But my belated realization about the emergent quality of spatial epidemiology has been really important and helpful for me. It made me realize that when I push into new areas - in life as much as research - that I probably have more of the tools I need than I realized in advance.

This means that you don't need to identify as a spatial analyst to be one. And if you want to think of yourself as one, you should, because ultimately it is the intention to engage with spatial relationships and geographic variation in a meaningful way that makes you a health geographer or spatial epidemiologist. This is likely true for many if not most scientific subfields[1], but this

[1]Though you probably need to be working near-ish to a field for it to happen by chance: there is little chance of me taking on the characteristics a particle physicist or chemical engineer by chance, but I wouldn't rule out lepidopterist or archaeologist entirely!

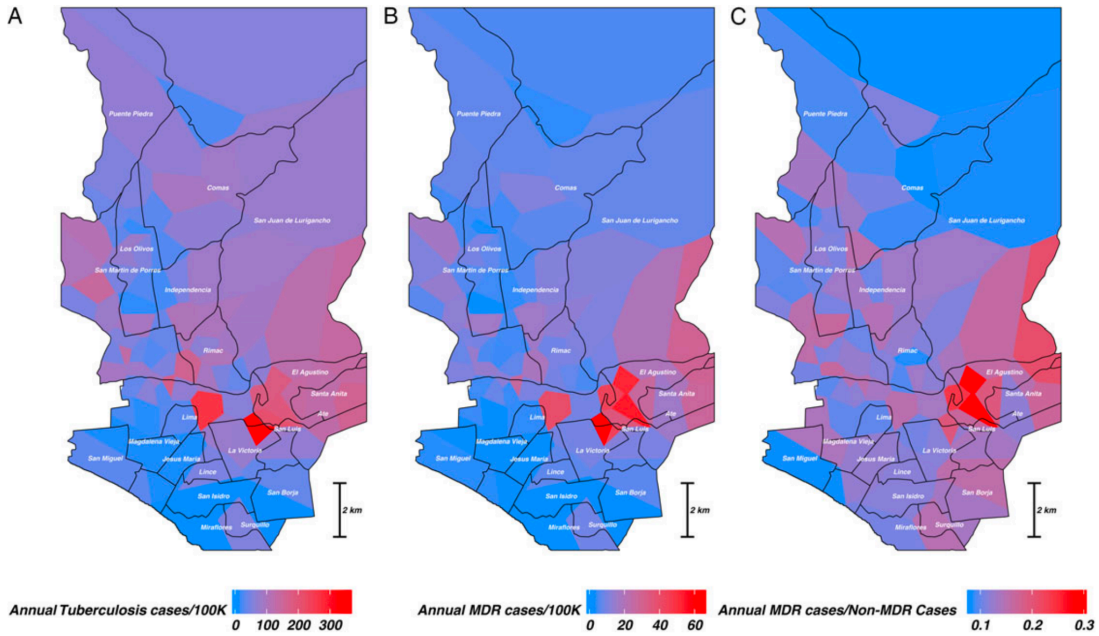

Figure 1. HC-level risks. Annual per-100 k rates of drug-sensitive and drug-resistant tuberculosis (*A*) and MDR tuberculosis (*B*), by HC catchment area. *C*, Ratio of the per-capita rate of MDR to non-MDR cases by HC. HC catchment areas are represented by polygons, with polygon fill color indicating the tuberculosis or MDR-tuberculosis rate in cases/100 K population. The boundaries of administrative districts of Lima are overlaid in black, and labeled in white**.** Abbreviations: HC, health center; MDR, multidrug-resistant.

Figure 3.6: The first map(s) I ever made, from (6), nearly 10 years after I started my research career.

one is mine and I'm glad I finally realized it!

References

Part II

Relationships

Part III

Space, time, and network

4 Space

4.1 Additional Resources

Please read the following two pieces that discuss key ideas about geospatial relatedness:

Miller HJ. Tobler's First Law and Spatial Analysis. *Annals of the Association of American Geographers*. 2004;94(2):284-289. doi:10.1111/j.1467-8306.2004.09402005.x

Goodchild MF. The Validity and Usefulness of Laws in Geographic Information Science and Geography. *Annals of the Association of American Geographers*. 2004;94(2):300-303. doi:10.1111/j.1467-8306.2004.09402008.x

5 Time

People

1. McElreath R. Statistical Rethinking: A Bayesian Course with Examples in R and STAN. 2nd edition. Boca Raton: Chapman and Hall/CRC; 2020.
2. Hilborn R, Mangel M. The Ecological Detective: Confronting Models with Data. 1st edition. Princeton, NJ: Princeton University Press; 1997.
3. Gelman A. Regression and Other Stories. 1st edition. Cambridge: Cambridge University Press; 2020.
4. Morris SE, Zelner JL, Fauquier DA, et al. Partially observed epidemics in wildlife hosts: Modelling an outbreak of dolphin morbillivirus in the northwestern Atlantic, June 2013–2014. *Journal of The Royal Society Interface* [electronic article]. 2015;12(112):20150676. (https://royalsocietypublishing.org/doi/10.1098/rsif.2015.0676). (Accessed December 15, 2019)
5. Thompson CN, Zelner JL, Nhu TDH, et al. The impact of environmental and climatic variation on the spatiotemporal trends of hospitalized pediatric diarrhea in Ho Chi Minh City, Vietnam. *Health & Place* [electronic article]. 2015;35:147–154. (https://linkinghub.elsevier.com/retrieve/pii/S1353829215001094). (Accessed December 15, 2019)
6. Zelner JL, Murray MB, Becerra MC, et al. Identifying Hotspots of Multidrug-Resistant Tuberculosis Transmission Using Spatial and Molecular Genetic Data. *Journal of Infectious Diseases* [electronic article]. 2016;213(2):287–294. (https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/jiv387). (Accessed December 15, 2019)