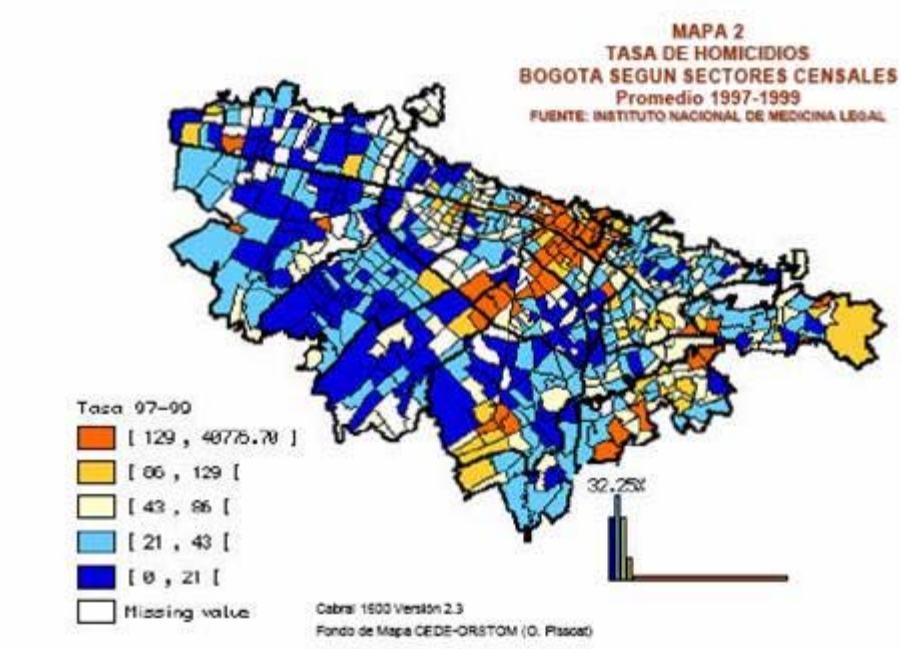
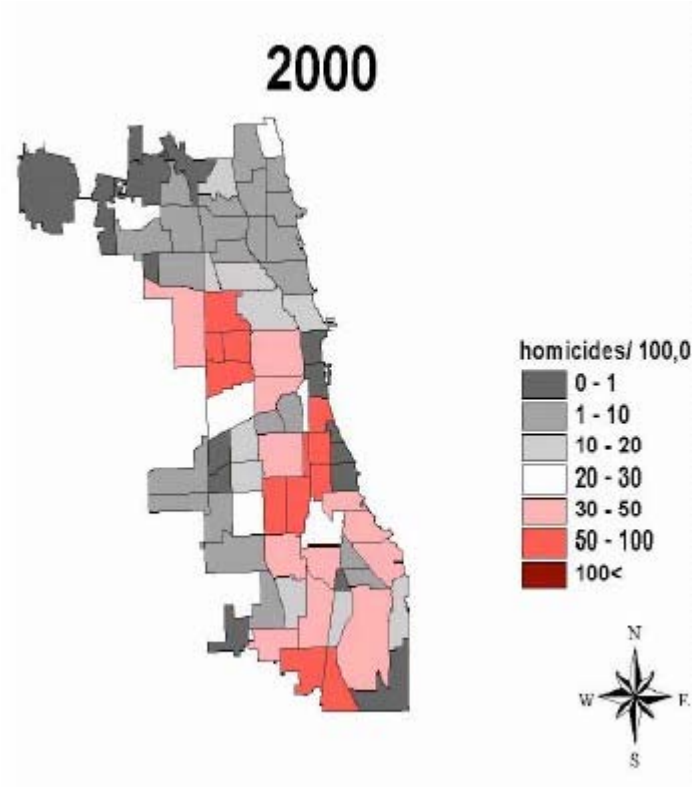
Potential Sampling Biases in Violence Surveys

Urban homicide rates can vary dramatically by location.

The map below gives homicide rates for census sectors in Bogotá, Colombia. It is not at all uncommon for areas with the lowest rates (dark blue) to border areas with the highest rates (oranges).



The same is true of every city for which I have seen evidence. Here is Chicago.

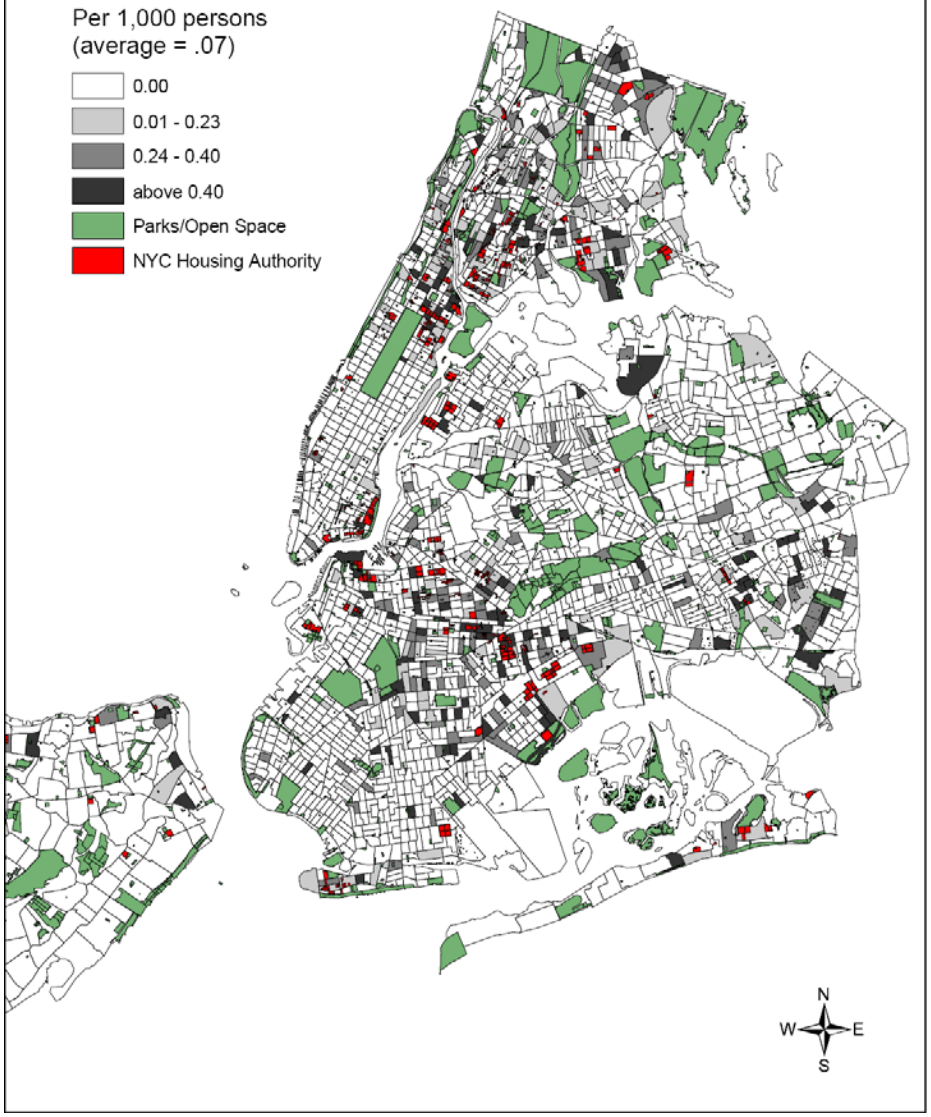


The Bogotá and Chicago maps contain a reasonable level of detail but do not penetrate down to the street level.

This homicide map of New York City still does not give a full street plan but it does have a fairly detailed scheme. Apply this microscope and apparently homogeneous areas are revealed as highly heterogeneous.

There are many places in the city where the homicide rate will change by a factor of more than 40 if you just walk a few blocks.

Figure 3a. Rate of Homicides, 2002



This map gives detailed information on a particular New York City neighborhood but still not a full street plan. Again, moving a few blocks can cause extraordinary variation.

Figure 3b. Rate of Homicides, 2002
Washington Heights and Inwood

What is the point?

If you try to use survey methods to measure urban homicide rates then minute details of the sampling scheme can matter quite a bit.

Suppose, for example, that you apply a basic variant or the "WHO-EPI" method to the Washington Heights and Inwood neighborhood. This might be done by:

1. Starting in the tiny little park in the middle of the neighborhood (the little green square).

2. Choose an angle between 0 and 360 degrees.

3. Travel in this direction to the edge of the neighborhood.

4. Suppose we can measure the homicide rate perfectly along this line. This is our estimate of the homicide rate in the neighborhood.

It is pretty clear from inspection of the diagram that this procedure will *underestimate* the homicide rate in the neighborhood.

A. *The method will place greater weight on areas near the center of the neighborhood than on areas at the extremities.*

For example, 0-homicide area immediately surrounding the little park gets included along any path from the center to the edge. The procedure discriminates particularly against areas along the long diagonal.

B. *The high-violence areas tend to be out at the extremities.*

Sudan mortality survey of Depoortere et al. published in the Lancet conducted field work within a number of camps for displaced people.

They found some very high design effects (as high as 11.3), even on deaths occurring *before* people arrived in the camps.

That is, families with similar pre-arrival mortality experiences tended to locate close to one another within the camps.

This makes sense. It could well be that people flee the same events together, arrive in camps together and are located by camp authorities nearby to one another.

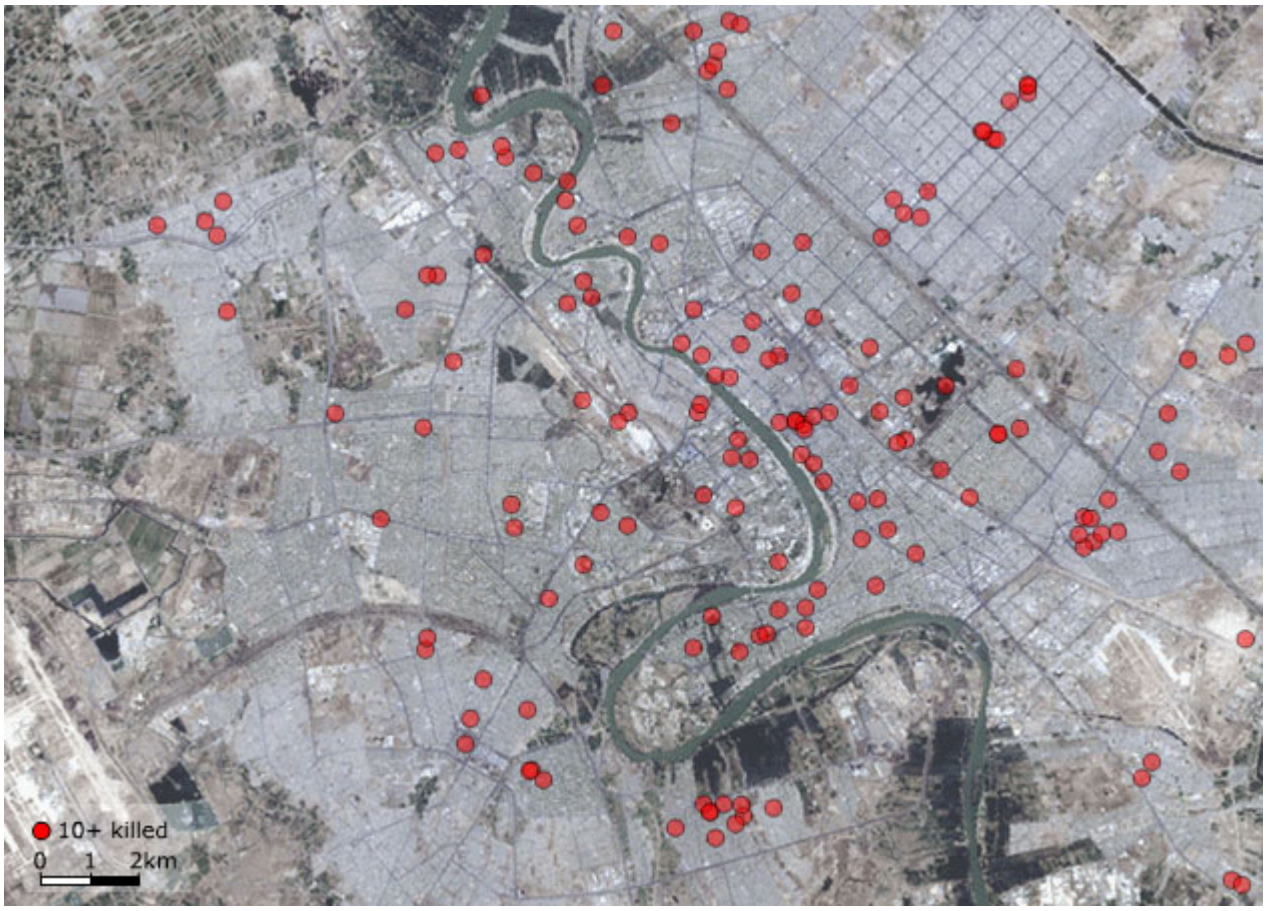
When such mechanisms are in play this can have strong implications for the properties of particular sampling procedures.

If there is reason to believe that mortality is higher (lower) at the extremities of a camp then procedures like WHO/EPI will tend to underestimate (overestimate) mortality rates.

It would be of great interest to see maps of within-camp variation.

Conflict violence has not been as well mapped as urban homicide but we do have some evidence. The following map is from the BBC website.

Baghdad: Mapping the violence
Attacks since May 2003 in which more than 10 people were killed.



Note that incidents of this size almost certainly cover over half of all deaths.

It is clear that there is a definite spatial pattern to these attacks.

Most of the dots are either on or near to streets that are big enough to stand out clearly on the map.

This makes sense.

1. Crowded markets, cafes restaurants and other attractions will be on such streets.

2. Military patrols focus on such streets. In fact, many military vehicles can only go down the larger streets.

3. Abductions and mass shootings will also tend to be on such streets. For example, Sunnis would not travel deep into Shiite territory, abduct some people and make a long drive to reach safe territory. Rather, they will make a quick foray in and out of enemy territory, perhaps just crossing over a main street that divides the two areas, just into a residential area.

The recent Burnham et al. Iraq study in the *Lancet* initiated every set of interviews from residential cross streets to main streets. The picture suggests that this could be a source of overestimation.

Here is a map pinpointing one particular bombing that killed 34 children who were gathering candy that had been tossed out by soldiers from a US Humvee.

This attack seems to have taken place precisely on a residential cross street to a main street.

There is another attack that actually appears in the dataset of the Lancet paper with 24 deaths out of the 66 actually occurring in the incident.

Baghdad market blast kills scores

A huge explosion has ripped through a busy Baghdad market, killing at least 66 people, officials say.

About 100 others were injured in the car bomb attack in Sadr City, a Shia area frequently targeted by insurgents.

The explosion left a scene of carnage and devastation, with the dead and injured lying amid the wreckage of cars, shops and market stalls.

The new government has been battling to improve Baghdad security, and last weekend unveiled a national unity plan.

But the attack was the worst incident in the capital for weeks.

Reports say the car bomb was detonated as a police patrol passed, causing both police and civilian casualties.

The bomb was clearly aimed at causing the maximum possible casualties, says the BBC's Jim Muir in Baghdad.

It was detonated at one of the busiest times of the day in a popular market in a densely-populated area.

Enlarge Image

In pictures: Baghdad blast

Biases in progression from household to household within a cluster:

Many published papers have a relatively detailed (although still inadequate) description of how they choose an initial household to begin a cluster of interviews.

But then they will write something particularly vague such as “We selected subsequent households by proximity, until completion of the cluster.” (Depoortere et al.)

Proximity could mean a variety of things:

1. Always move to the household nearest to the one where you have just completed an interview. This would often imply crossing a street or an alley. In a refugee camp it could mean fanning out in concentric circles. Correct application requires some fairly sophisticated measurement.

2. Stay on one side of a street (if you are interviewing in a street environment) and proceed in a single pre-defined direction, e.g., east, until you are finished. At corners you just cross the street and keep going in the same direction.

3. (An alternative to 2) Change directions when you reach a corner. An obvious approach is to turn at a 90 degree angle (depending on street layout).

4. Proceed along a line, as in 1, but in both directions, e.g., both east and west.

The point is:

1. *Proximity* alone is *ambiguous*.

2. Where there is *ambiguity* there is *discretion* for field teams.

3. When there is *discretion* there is opportunity for *biases* to come into play.

Consider this scenario

A field team works its way along the street conducting interviews.

Their field protocols call for them to proceed by proximity.

They reach a corner.

They look left and see a normal-looking block.

They look right and some bomb-out houses and bullet holes in some walls.

They can proceed either to the left or to the right and be faithful to the vague notion of proximity which they have been asked to follow.

Which way do they turn?

If interviewers have biases this is a perfect opportunity to express them by turning toward the battle-scared block (or, less common I believe, away if they wish to suppress deaths).

Even with the best of intentions there will be a normal human urge to go where the deaths are likely to be and to bear witness.

Many people will feel a strong moral urge to gravitate towards likely deaths that will trump abstract notions of proper measurement.

Such behavior will cause upward bias.

Discretion in field protocols enable this bias to express itself even by well-meaning interviewers.

Solid field protocols would specify clearly where the team should go and it the team travels in the wrong direction then they would be cheating, plain and simple.

Variations on the theme of field discretion opening the door to bias:

1. People approach the field team asking to be interviewed or offering to lead the team to important households that should be interviewed.

2. People who have been recently displaced from a region also covered by a survey turning up in the sample in a different region. (Unless such migrants are systematically incorporated into a survey such households will need to be excluded.)