Measuring exposure to location-based risk factors during daily activities in urban landscapes

Douglas Wiebe, PhD
Department of Biostatistics & Epidemiology
University of Pennsylvania

Consortium for Education and Social Science Research
Indiana University
Workshop in Methods

March 4, 2011
Map used by John Snow in describing the Broad Street pump cholera outbreak of 1854
Epidemiologic Triad
Space-Time Adolescent Risk Study

Paul Allison PhD, Elijah Anderson PhD*, Charles Branas PhD, Wensheng Guo PhD, Judd Hollander MD, Michael Nance MD, C. William Schwab MD, Therese Richmond PhD RN, C. Dana Tomlin PhD, Douglas Wiebe PhD

University of Pennsylvania
*Yale University

National Institute on Alcohol Abuse and Alcoholism, National Institute of Child Health & Human Development (R01AA014944)

Graphic provided by The HELP Network, Chicago
Alcohol outlets: n=1700
STARS is a “population-based case-control study”

<table>
<thead>
<tr>
<th>Type of study</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>Studies preventions and treatments for diseases; investigator actively manipulates which groups receive the agent under study.</td>
</tr>
<tr>
<td>Cohort</td>
<td>Typically examines multiple health effects of an exposure; subjects are defined according to their exposure levels and followed for disease occurrence.</td>
</tr>
<tr>
<td>Case–control</td>
<td>Typically examines multiple exposures in relation to a disease; subjects are defined as cases and controls, and exposure histories are compared.</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>Examines relationship between exposure and disease prevalence in a defined population at a single point in time.</td>
</tr>
<tr>
<td>Ecological</td>
<td>Examines relationship between exposure and disease with population-level rather than individual-level data.</td>
</tr>
</tbody>
</table>
Design of cohort study vs case-control study

Cohort Studies → Factor

<table>
<thead>
<tr>
<th>Present (exposed)</th>
<th>What % got disease (risk E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(follow forward in time)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Absent (not exposed)</th>
<th>What % got disease (risk ( \overline{E} ))</th>
</tr>
</thead>
</table>
Design of cohort study vs case-control study

Case-Control Studies

Disease

Present (cases) Absent (controls)
(look backward in time)

What % had been exposed What % had been exposed
Design of cohort study vs case-control study

Cohort study

RR = \frac{A}{A+B} \frac{C}{C+D}

Case-control study

OR = \frac{AD}{BC}

Case-Control Studies

Disease

Present (cases) Absent (controls)

A B

C D

Cohort Studies
Factor

Present (exposed)

Absent (not exposed)
Table 4. Example 24-h recall diary containing beginning and ending times, activity, location, presence of a smoker, and time spent for 22 microenvironments visited on the diary day.

<table>
<thead>
<tr>
<th>Microenvironment number</th>
<th>Starting time</th>
<th>Ending time</th>
<th>Summary</th>
<th>Detailed activity</th>
<th>Simplified activity</th>
<th>Detailed location</th>
<th>Simplified location</th>
<th>Smoker? (1 = Yes)</th>
<th>Time spent (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>00:00</td>
<td>01:45</td>
<td>At night club</td>
<td>77</td>
<td>0</td>
<td>405</td>
<td>90</td>
<td>1</td>
<td>105</td>
</tr>
<tr>
<td>2</td>
<td>01:45</td>
<td>02:00</td>
<td>Traveled home after night club</td>
<td>79</td>
<td>0</td>
<td>301</td>
<td>30</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>02:00</td>
<td>11:00</td>
<td>Sleeping or napping</td>
<td>45</td>
<td>0</td>
<td>105</td>
<td>10</td>
<td>0</td>
<td>540</td>
</tr>
<tr>
<td>4</td>
<td>11:00</td>
<td>11:05</td>
<td>Brushed teeth</td>
<td>44</td>
<td>40</td>
<td>104</td>
<td>10</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>11:05</td>
<td>11:15</td>
<td>Preparing meals or snacks</td>
<td>10</td>
<td>10</td>
<td>101</td>
<td>10</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>11:15</td>
<td>11:25</td>
<td>Eating meals or snacks</td>
<td>43</td>
<td>70</td>
<td>102</td>
<td>10</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>11:25</td>
<td>11:30</td>
<td>Dressing or personal grooming</td>
<td>47</td>
<td>0</td>
<td>102</td>
<td>10</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>11:30</td>
<td>11:37</td>
<td>Traveling to play football</td>
<td>89</td>
<td>0</td>
<td>306</td>
<td>40</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>11:37</td>
<td>13:37</td>
<td>Playing flag football</td>
<td>80</td>
<td>60</td>
<td>507</td>
<td>50</td>
<td>0</td>
<td>120</td>
</tr>
<tr>
<td>10</td>
<td>13:37</td>
<td>13:44</td>
<td>Traveling to home</td>
<td>79</td>
<td>0</td>
<td>306</td>
<td>40</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>11</td>
<td>13:44</td>
<td>13:54</td>
<td>Preparing meals or snacks</td>
<td>10</td>
<td>10</td>
<td>201</td>
<td>10</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>12</td>
<td>13:54</td>
<td>13:57</td>
<td>Traveling to bar</td>
<td>79</td>
<td>0</td>
<td>301</td>
<td>30</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td><strong>13</strong></td>
<td><strong>13:57</strong></td>
<td><strong>15:30</strong></td>
<td><strong>At bar</strong></td>
<td><strong>77</strong></td>
<td><strong>0</strong></td>
<td><strong>405</strong></td>
<td><strong>90</strong></td>
<td><strong>1</strong></td>
<td><strong>93</strong></td>
</tr>
<tr>
<td>14</td>
<td>15:30</td>
<td>15:33</td>
<td>Traveling from bar</td>
<td>79</td>
<td>0</td>
<td>301</td>
<td>30</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>15:33</td>
<td>16:30</td>
<td>Watching TV</td>
<td>91</td>
<td>0</td>
<td>102</td>
<td>10</td>
<td>0</td>
<td>57</td>
</tr>
<tr>
<td>16</td>
<td>16:30</td>
<td>17:00</td>
<td>Bathing or showering</td>
<td>40</td>
<td>40</td>
<td>104</td>
<td>10</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>17</td>
<td>17:00</td>
<td>19:00</td>
<td>Watching TV</td>
<td>91</td>
<td>0</td>
<td>102</td>
<td>10</td>
<td>0</td>
<td>120</td>
</tr>
<tr>
<td>18</td>
<td>19:00</td>
<td>19:10</td>
<td>Traveling to shopping</td>
<td>39</td>
<td>0</td>
<td>301</td>
<td>30</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>19</td>
<td>19:10</td>
<td>19:25</td>
<td>Shopping for food</td>
<td>30</td>
<td>0</td>
<td>414</td>
<td>90</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>20</td>
<td>19:25</td>
<td>19:35</td>
<td>Travel related to shopping for food</td>
<td>39</td>
<td>0</td>
<td>301</td>
<td>30</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>21</td>
<td>19:35</td>
<td>21:00</td>
<td>Watching TV</td>
<td>91</td>
<td>0</td>
<td>102</td>
<td>10</td>
<td>0</td>
<td>85</td>
</tr>
<tr>
<td>22</td>
<td>21:00</td>
<td>24:00</td>
<td>Studying</td>
<td>54</td>
<td>0</td>
<td>102</td>
<td>10</td>
<td>0</td>
<td>180</td>
</tr>
</tbody>
</table>

The respondent, whose diary is shown in this table, was a Hispanic male from Connecticut between the ages of 18 and 24 who was interviewed on a weekend in the fall. See the Sample and Data characteristics section for a description of the simplified (i.e., recoded) locations and activities.
Proportion of time spent in each of six locations

NHAPS - Nation, Percentage Time Spent
Total n = 9,196

- In a residence (68.7%)
- Indoors (86.9%)
- Outdoors (7.6%)
- In a vehicle (5.5%)
- Other indoor location (11%)
- Office-factory (5.4%)
- Bar-restaurant (1.8%)

These are “activity pattern data”

Klepeis et al. J Expo Analysis & Environ Epi, 2001
Radon is a colorless, odorless, naturally occurring, radioactive noble gas that is formed from the decay of radium.

Radon is a significant contaminant that affects indoor air quality worldwide. Radon gas from natural sources can accumulate in buildings, especially in confined areas such as the basement and reportedly causes 21,000 lung cancer deaths per year in the US.

Radon is the second most frequent cause of lung cancer, after cigarette smoking, and radon-induced lung cancer is thought to be the 6th leading cause of cancer death overall.

Radon can be found in some spring waters and hot springs.

Gatrell & Loytonen, GIS & Health 1998, p.106
Recruitment

Case subjects: HUP and CHOP
- Screening by Academic Associates
- Interviewing by full-time project staff
- Interview takes place in ER, on hospital ward, home, or research office

Control subjects: community
- Screening via RDD (random digit dialing)
- Interviewing by full-time project staff
- Interview takes place at home or research office
How are you getting around? Here are some examples. Others?

What are you doing? Anything else?

How safe do you feel?
On a scale of 1-10, how safe do you feel?
10 FEELING VERY SAFE
1 FEELING VERY UNSAFE

Are any of these things involved? Anything else?

Who are you with? Family, Friends, Girlfriend, Boyfriend, Someone you don’t like, anyone else?
Census tracts: n=381
Time-varying exposures and covariates

Figure 1. Dichotomous variable

Figure 2. Continuous variable

Figure 3. Gravity measure

\[ \text{Alcohol store}_i = \sum_j S_{ij} / \min(d_{ij}, c) \]
\[ = (l_1/60_{11}) + (l_2/250_{12}) = 2.07 \times 10^{-2} \]
\[ = (l_1/300_{11}) + (l_2/350_{12}) = 0.62 \times 10^{-2} \]

Computing a gravity measure of exposure to alcohol stores on 4 city blocks, for subject 1 at time 1 and time 2.
Neighborhoods: n=69
Figure 1. Ten subjects' hand-drawn depictions of the land area they consider to be their neighborhoods, shown in conjunction with the location of their home.

Area(s) refers to approximate land area of what each subject hand-drew on a map to represent their neighborhood.

X = Location of subject's residence.
Figure 2. Paths of 10 subjects' 24-hour daily activities, shown in conjunction with their hand-drawn neighborhood depletion and home location.

Areas 1-3 = <0.1 sq mi
GLDH = 0.8 ml
Tot dis = 2.8 ml

Areas 4-6 = 0.2-0.7 sq mi
GLDH = 4.1-4.7 ml
Tot dis = 10.1-12.5 ml

Areas 7-9 = >0.7 sq mi
GLDH = 11.8-19.7 ml
Tot dis = 16.5-25.9 ml

Legend:
- X = Location of subject's residence.
- GLDH = Greatest linear distance from home.
- Tot dis = Route total distance.
Characteristics of one-day activity paths of 15-19 year-old subjects (n=55)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Min</th>
<th>Max</th>
<th>Mean (SD)</th>
<th>Median (25%, 75%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (hrs) spent outside home during one-day reporting period</td>
<td>0.0</td>
<td>23.5</td>
<td>7.9 (5.6)</td>
<td>8.3 (3.5, 10.9)</td>
</tr>
<tr>
<td>Greatest linear distance from home (mi)</td>
<td>0.0</td>
<td>7.2</td>
<td>1.4 (1.7)</td>
<td>0.9 (0.2, 2.4)</td>
</tr>
<tr>
<td>Distance travelled (mi)</td>
<td>0.0</td>
<td>20.6</td>
<td>4.5 (4.7)</td>
<td>3.2 (0.7, 9.1)</td>
</tr>
<tr>
<td>Time (hrs) spent outside census tract of residence during one-day reporting period</td>
<td>0.0</td>
<td>19.8</td>
<td>6.3 (5.5)</td>
<td>7.0 (0.0, 9.4)</td>
</tr>
<tr>
<td>Proportion of outside-the-home time that was spent outside census tract of residence (%)</td>
<td>0.0</td>
<td>99.4</td>
<td>71.0 (36.5)</td>
<td>91.5 (56.9, 96.7)</td>
</tr>
<tr>
<td>Number of census tracts intersected by subject's one-day activity path</td>
<td>1</td>
<td>34</td>
<td>8.1 (8.5)</td>
<td>6 (1, 10)</td>
</tr>
</tbody>
</table>

25% and 75% denote the twenty-fifth and seventy-fifth percentiles, respectively.

In most instances (87.8%), it was one of four types of locations that constituted the place along a subject’s path that was the farthest point (ie, linear distance) from their home: school, work, places of recreation, and food stores and restaurants.
Median values of the greatest linear distances traveled from home:

**1.1 mi on weekdays vs 0.4 mi on weekends**
(Mann-Whitney z=1.83, p=0.07)

- subjects’ activities generally involved staying closer to home on weekends

Cumulative distances travelled were generally shorter on weekends also (medians):

**4.2 mi on weekdays vs 1.2 mi on weekends**
(Mann-Whitney z=1.89, p=0.09)
Various measures of exposure to alcohol outlets for 15-19 year-old subjects (n=55)

<table>
<thead>
<tr>
<th>Alcohol outlet prevalence in subject's census tract of residence</th>
<th>Per 1000 residents</th>
<th>Per road kilometer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>46.4</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>5.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Alcohol outlets contacted during one-day activities</th>
<th>Walked within 18 meters</th>
<th>Walked within 200 meters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>120</td>
</tr>
</tbody>
</table>

25% and 75% denote the twenty-fifth and seventy-fifth percentiles, respectively.
Relation between alcohol outlet prevalence in subjects’ census tracts of residence and alcohol outlets contacted during one-day activities

Spearman’s rho = 0.06
p = 0.66
Experiences in their neighborhoods: 13-17 controls vs. case subjects*

I've heard gunshots
I've seen people using or selling drugs
I often see drunk people on the street
I've seen someone get stabbed
I've seen someone pull a gun on somebody
I've seen someone get shot

% of respondents

*p<0.05

*Partial data; preliminary
Interviewer’s heart rate (actual and fitted) during a 150-minute interview trip from campus to a subject’s home and back

Note: Fitted results were derived from an autoregressive integrated moving average (ARIMA) (1,0,0) model ($\phi=0.2$, $p=0.05$; constant=$91.8$, $p=0.001$) that produced white noise residuals ($Q=27.3$ at 24 lags).

† Gradual, permanent heart rate change modeled with a first order transfer function applied to a step variable. Denominator was constrained to 0.7.

‡ Abrupt, temporary heart rate change modeled with a first order transfer function applied to a pulse variable. Denominator was constrained to 0.7.

* $p<0.05$
KEY POINTS

The notion that exposure levels can vary in meaningful ways over the course of daily activities has been noted in some research fields, but may remain under-appreciated in large-scale studies of neighborhoods and health.

These results illustrate how individuals’ daily activities frequently occur in locations away from their homes, cross administrative boundaries, and likely result in exposure to environmental exposures at very different levels than at their home.

Measuring exposure should thus be carefully considered in light of the exposure-disease relation under study and the induction period of the exposure of interest.

CONCLUSION

In studies of environmental exposures that are encountered outside the home and that vary in etiologically meaningful ways over the course of daily activities, classifying subjects as exposed based solely on the prevalence of the exposure in the administrative geographic unit of their residence (e.g. a Census tract or ZIP code) will likely result in exposure misclassification.

Ask, Where do subjects spend time? And how much time do they spend there? Both will be key to better understanding “neighborhood.”