Overview

- Background
- Study Purpose
- Methods
- Results
- Strengths and Limitations
- Implications and Future Research
Background
Smoking and Tobacco Advertising

- Smoking is the leading cause of preventable death in U.S., accounting for approximately 440,000 deaths annually
  - Mokdad, et al, 2004

- Tobacco advertising influences smoking-related attitudes and behaviors
  - NCI Monograph 19, Role of Media in Promoting and Reducing Tobacco Use, 2008
Retail is the most important advertising channel for the tobacco industry.
Impact of Retail Tobacco Advertising

- Youth who are exposed to point-of-sale (POS) tobacco marketing are more likely to experiment with smoking and to become smokers.

- For adults, POS tobacco marketing may act as cue to smoke, stimulate unplanned purchases, and encourage relapse among recent quitters
  Carter, Mills, Donovan (2009); Wakefield, Germain, Durkin (2008)
Source: Counter Tobacco
Ongoing Surveillance Needs

- New products (electronic cigarettes) introduced
- 2009 Family Smoking Prevention and Tobacco Control Act gives states and local governments the authority to regulate the manner, place, and time of advertising
Retail Audits: Traditional Approach

- Small group of trained data collectors travel to stores
- Collect data inconspicuously
- 3-6 page survey, 15 mins to complete
- N=hundreds to thousands
- Weeks to months
- Costs for training, travel, data collection, management
- Time consuming, cost-prohibitive
Crowdsourcing: Alternative Approach?

- Divide audit into smaller, scalable components
- Leverage via the Internet, mobile workforce to conduct audits in their neighborhoods
- More rapid, sustainable surveillance system?
- Gaps
  - limited application for retail data collection
  - Limited info on quality of crowdsourced data
Purpose

- Explore the feasibility of crowdsourcing retail audits by comparing the quality of data collected by untrained crowdsourced workers to trained data collectors
Methods

- Annual tobacco retail audit for Bureau of Tobacco Free Florida, trained surveyors conducted in-store audits in Miami and Tampa from 8/31/2012 to 9/19/2012
- Used crowdsourcing application Gigwalk, mobile workforce of 325,000 workers
- Crowdsourced workers collected data at the same stores during the same time period
How Gigwalk Works

1. Post Job ("Gig")

2. Gigs appear on mobile apps

3. Workers accept and conduct gigs

4. Workers upload data and are paid

Here's how it works: You have products to check or information to capture at thousands of locations, like local businesses or retail stores.
Methods (cont)

- Posted 194 gigs on Gigwalk, to workers in the Miami & Tampa areas - “cigarette” (99) and “smokeless” (95)
- Measures
  - Products for sale: e-cigarettes (flavored), snus, dissolvables
  - Promotions for Marlboro cigarettes: sales price, coupon, multipack discount, free pack with purchase, mail in rebate
  - Presence of cigarette ads: exterior, interior of store
  - Photo: e-cigarette product, ads posted on store exterior
- Crowdsourced workers uploaded data and photos via Gigwalk mobile app and were paid $7
- Analysis: Percent agreement and kappa statistics
Results

- Of 194 jobs posted online, 78.3% were initiated and 72.2% were completed
- 25 crowdsourced workers
- # of audits per worker: avg = 3, range = 1 to 25 stores
- Time to complete audit and submit data: median = 18.1 hours within posting the job online
- Comparison sample = 109 completed stores
  - 55 stores cigarette audit, 54 stores smokeless audit
### Results (cont)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Crowdsourced Untrained Worker</th>
<th>Trained Data Collector</th>
<th>% Agreement, Kappa statistic, p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Products for Sale</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-cigarettes</td>
<td>42.50%</td>
<td>33.30%</td>
<td>79.6%, k=0.57, p=0.000</td>
</tr>
<tr>
<td>E-cigarettes (flavored)</td>
<td>37.00%</td>
<td>20.40%</td>
<td>72.2%, k=0.34, p=0.003</td>
</tr>
<tr>
<td>Snus</td>
<td>22.20%</td>
<td>22.20%</td>
<td>85.2%, k=0.57, p=0.000</td>
</tr>
<tr>
<td>Dissolvables</td>
<td>7.40%</td>
<td>0.00%</td>
<td>92.6%, k=0.00, p=n/a</td>
</tr>
</tbody>
</table>

**Kappa strength of agreement**
- 0 - 0.30: poor to slight
- 0.31 – 0.6: moderate
- 0.61 – 1.0: strong

*Note: kappa limitations - does not account for bias and prevalence so kappa is low even when agreement is high*
<table>
<thead>
<tr>
<th>Measure</th>
<th>Crowdsourced Untrained Worker</th>
<th>Trained Data Collector</th>
<th>% Agreement, Kappa statistic, p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tobacco Advertising</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence of exterior cigarette ad</td>
<td>50.90%</td>
<td>54.60%</td>
<td>85.5%, k=0.71, p=0.000</td>
</tr>
<tr>
<td>Presence of interior cigarette ad</td>
<td>3.70%</td>
<td>11.10%</td>
<td>88.9%, k=0.21, p=0.037</td>
</tr>
<tr>
<td>Promotions for Marlboro Cigarettes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Sale/special/reduced” price listed</td>
<td>29.10%</td>
<td>3.60%</td>
<td>67.3%, k=-0.07, p=0.822</td>
</tr>
<tr>
<td>Multi-pack discount</td>
<td>16.40%</td>
<td>10.90%</td>
<td>87.3%, k=0.46, p=0.000</td>
</tr>
<tr>
<td>Mail-in rebate</td>
<td>0.00%</td>
<td>0.00%</td>
<td>100%, k=n/a, p=n/a</td>
</tr>
<tr>
<td>Coupon attached to pack</td>
<td>3.60%</td>
<td>0.00%</td>
<td>96.4%, k=0.00, p=n/a</td>
</tr>
<tr>
<td>Free pack(s) with purchase</td>
<td>1.80%</td>
<td>0.00%</td>
<td>98.2%, k=0.00, p=n/a</td>
</tr>
</tbody>
</table>
Summary + Discussion

- High agreement between crowdsourced worker and trained data collectors for most measures.
- Low agreement for more nuanced measures like presence of sales price, which may require training.
- For some measures, crowdsourced workers may have provided more accurate data, e.g. e-cigarette availability confirmed by photos.
Crowdsourced
23 Stores

Trained
18 stores

Validated: 21 stores
Could not validate: 2 stores
Strengths:
- First study that compared quality of retail audit data collected by crowdsourced workers vs. trained data collectors
- Same measures collected from same stores

Limitations:
- Data collected during same week, but it’s possible store environment could have changed depending on when crowdsourced worker vs. trained data collector visited the store
- Unable to directly compare costs
- Only examined small number of measures, results may not generalize to all tobacco retail audit measures

<table>
<thead>
<tr>
<th>Crowdsourced</th>
<th>Trained</th>
</tr>
</thead>
<tbody>
<tr>
<td>$7</td>
<td>$60-70 (other costs)</td>
</tr>
<tr>
<td>6-7 measures</td>
<td>50-60 measures</td>
</tr>
<tr>
<td>5 mins?</td>
<td>avg:15 mins</td>
</tr>
</tbody>
</table>
Implications + Future Research

- Crowdsourcing may be a viable option for collecting retail audit data
- Photos and geolocation data can be used to validate crowdsourced data
- Local residents’ familiarity with neighborhood could facilitate data collection (e.g. asking clerks directly about product availability instead of collecting data inconspicuously)
- Future studies should examine which measures are most amenable to crowdsourcing and the relative cost of crowdsourcing compared to using trained data collectors
More Information

Annice Kim, PhD
Senior Social Scientist
919.316.3972
akim@rti.org