Use of Google Street View to Assess Environmental Contributions to Pedestrian Injury

Stephen J. Mooney, MS, Charles J. DiMaggio, PhD, Gina S. Lovasi, PhD, Kathryn M. Neckerman, PhD, Michael D. M. Bader, PhD, Julien O. Teitler, PhD, Daniel M. Sheehan, MA, Darby W. Jack, PhD, and Andrew G. Rundle, DrPH, MPH

Objectives. To demonstrate an information technology–based approach to assess characteristics of streets and intersections associated with injuries that is less costly and time-consuming than location-based studies of pedestrian injury.

Methods. We used imagery captured by Google Street View from 2007 to 2011 to assess 9 characteristics of 532 intersections within New York City. We controlled for estimated pedestrian count and estimated the relation between intersections' characteristics and frequency of injurious collisions.

Results. The count of pedestrian injuries at intersections was associated with the presence of marked crosswalks (80% increase; 95% confidence interval [CI] = 2%, 218%), pedestrian signals (156% increase; 95% CI = 69%, 259%), nearby billboards (42% increase; 95% CI = 7%, 90%), and bus stops (120% increase; 95% CI = 51%, 220%). Injury incidence per pedestrian was lower at intersections with higher estimated pedestrian volumes.

Conclusions. Consistent with in-person study observations, the information-technology approach found traffic islands, visual advertising, bus stops, and crosswalk infrastructures to be associated with elevated counts of pedestrian injury in New York City. Virtual site visits for pedestrian injury control studies are a viable and informative methodology. (*Am J Public Health.* 2016;106:462–469. doi:10.2105/AJPH.2015.302978)

See also Galea and Vaughan, p. 394.

In 2013, an estimated 70 000 pedestrians were injured or killed by motor vehicles in the United States.[1] Pedestrian deaths as a proportion of overall traffic fatalities have grown from 11% to 14% in the past decade nationally,[2] and in New York City more pedestrians than vehicle occupants have been killed by motor vehicles each year since at least 1910.[3] On a trip-by-trip basis, a pedestrian is 50% more likely to be killed than a motor vehicle occupant.[4] Pedestrian safety is not only vital for public health directly through reduced traffic-related morbidity and mortality, but also indirectly as the perception of increased safety from traffic encourages outdoor physical activity, with consequent mental and physical health benefits.[5,6] Investments in pedestrian safety infrastructure may be a particularly cost-effective way to improve population health.[7] Recently, several major cities, including New York City, have developed high-profile plans to improve pedestrian safety citywide.[8] New York City alone has installed 1500 pedestrian signals and reengineered dozens of roads and intersections.[9]

Although recent studies have suggested that road and pedestrian environment modifications such as improving lighting, adding speed bumps, or maintaining pavement markings can substantially improve pedestrian safety,[10–13] such studies have not been as widely replicated, in part because of methodological and logistical challenges of conducting pedestrian environment studies. One 2003 systematic review of literature regarding engineering measures designed to reduce pedestrian–motor vehicle crashes that considered both before and after interventions and location-based case–control designs found at most 3 studies per intervention type and only 1 study for most intervention types.[11] The ability to replicate findings is essential to the scientific method and the lack of replication studies undermines weight of evidence and meta-analytical approaches to determining evidence-based best practices.

The large burden of injury coupled with the sparse empirical literature justifies more research into risk factors for pedestrian injury and into the effectiveness of interventions. Such studies have, however, typically entailed costly data-collection procedures that require researchers to visit and code every intersection included in the study. Recently, several research teams have validated the use of Google Street View to conduct "virtual" neighborhood audits that obviate the need for field teams to conduct in-person audits.[14–18] However, to the best of our knowledge, such approaches have not been used to assess risk factors for pedestrian injury. In this article, we demonstrate an investigation of environmental contribution to injuries by using information technology rather than site visits.

METHODS

We drew a systematic sample of 532 intersections from New York City. The sample comprised an approximately 2-kilometer by 2-kilometer grid across the entire city with oversamples in poor and the most densely

ABOUT THE AUTHORS

Stephen J. Mooney, Charles J. DiMaggio, Gina S. Lovasi, Daniel M. Sheehan, and Andrew G. Rundle are with Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY. Kathryn M. Neckerman is with Columbia Population Research Center, Columbia University. Michael D. M. Bader is with Department of Sociology, American University, Washington, DC. Julien O. Teitler is with School of Social Work, Columbia University. Darby W. Jack is with Department of Environmental Health Sciences, Mailman School of Public Health.

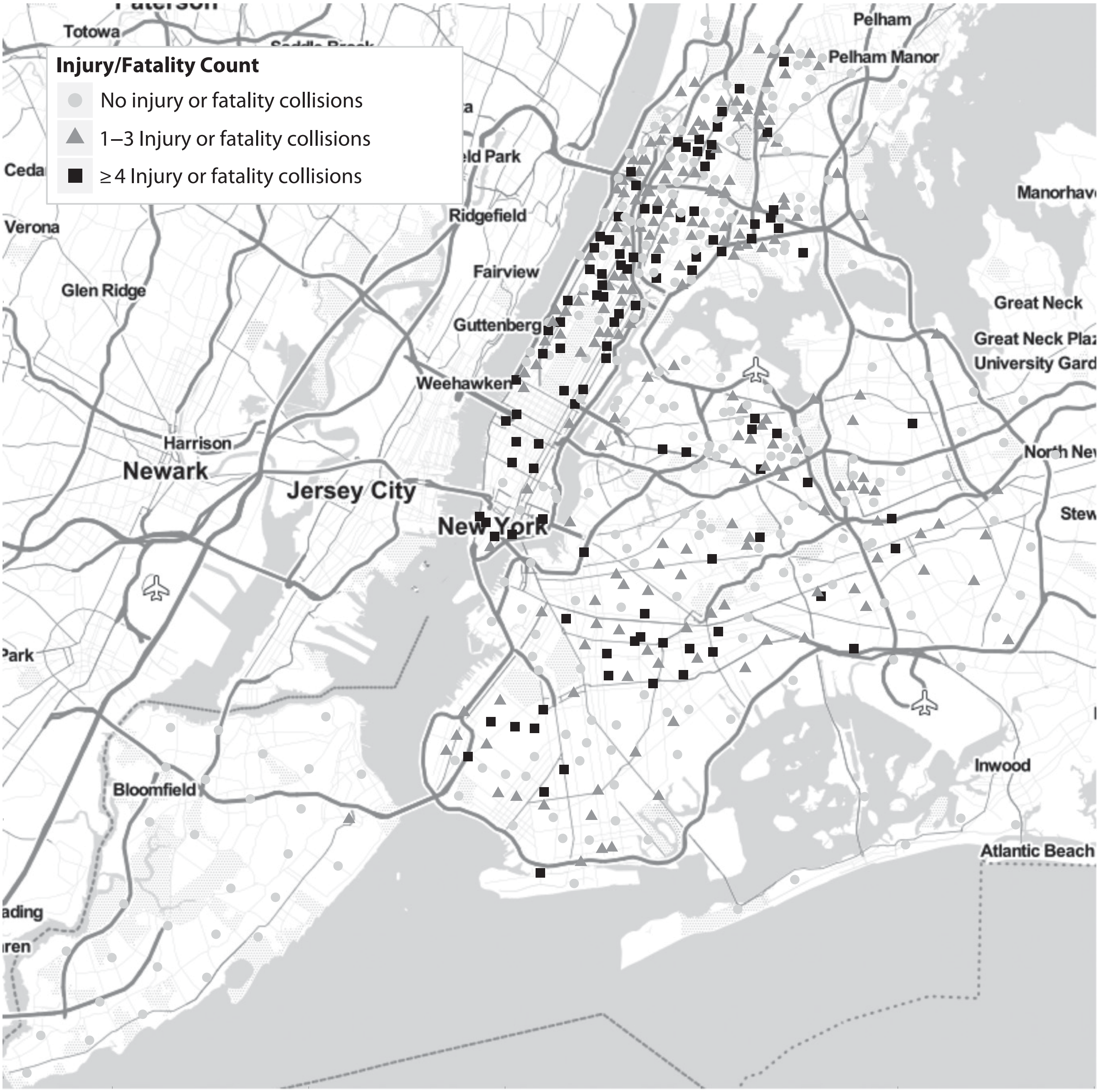
Correspondence should be sent to Andrew Rundle, DrPH, MPH, 722 W 168th St, Room 727, Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY 10032 (e-mail: agr3@cumc.columbia.edu). Reprints can be ordered at http://www.ajph.org by clicking the "Reprints" link.

This article was accepted October 31, 2015.

doi: 10.2105/AJPH.2015.302978

populated areas. We used an autosampler (described in detail previously[19]), to map these locations to block faces with Google Street View in New York City (Figure 1).

We assessed each intersection for pedestrian environments with the Computer Assisted Neighborhood Visual Assessment System (CANVAS), an efficient, validated tool for virtual street audits.[20] CANVAS includes a management interface, allowing researchers to develop samples, train auditors, monitor progress, and measure reliability, as well as an auditor interface allowing virtual street auditors to manipulate a Google Street View to view intersections from varying perspectives and collect data efficiently. Five trained virtual street auditors collected data between July 2012 and March 2013. Auditors assessed 64 previously validated[20] items to assess urban context, though not all items were directly relevant to pedestrian safety. Audits required approximately 10 minutes per intersection. Google initially captured the Street View imagery used in this study between August 2007 and October 2011. We adapted items to assess intersection characteristics from the Irvine–Minnesota



Source. Open Street Map (http://www.openstreetmap.org).

FIGURE 1—Injury or Fatality Counts at Intersections Virtually Visited: New York City, 2007–2011

Inventory[21] and the Pedestrian Environment Data Scan[22] as described by Bader et al.[20] To assess interrater reliability, we randomly selected 5% of intersections to be audited multiple times. We selected 1 rating at random from intersections in the reliability sample for final analysis.

Intersection characteristics available for this study were presence of crosswalks (none, some, all, or not applicable); presence of curb cuts (yes or no); presence of visible billboards (yes or no); presence of sidewalk (complete, incomplete, or none); condition of sidewalk where present (poor, fair, or good); condition of road (poor, fair, good, or under repair); presence of pedestrian signal (yes or no); presence of any type of traffic or pedestrian refuge (yes or no); presence of traffic-calming devices such as curb extensions, chicanes, or speed bumps (yes or no); and presence of bus stop (yes or no; Table A, available as a supplement to the online version of this article at http://www.ajph.org).

We derived collision data from 2 public-use databases (CrashStat and CrashMapper) comprising all collisions within New York City to which police were dispatched. The New York State Department of Transportation initially compiled the data from police records; Transportation Alternatives, an advocacy organization, worked to make the data public. These data have been used in previous studies to investigate determinants of collisions involving pedestrians[9] and other health outcomes.[23,24] We focused on pedestrian–vehicle collisions occurring between 2007 and 2011, years corresponding to the date range of imagery available on Google Street View at the time the audit was conducted. To account for geographically large intersections (for example, Queens Boulevard is 69 meters across) and variation between the geocoding tools used in creating CrashStat and Google Maps' geocoder, we considered collisions to have occurred at an intersection if their recorded geographic coordinates were within 60 meters of the center of that intersection, though, as described in the "Sensitivity Analysis" subsection of Methods, we also performed a sensitivity analysis requiring collisions to be within 30 meters. Figure 1 indicates the number of collisions occurring at each of the selected intersections on a map of New York City.

Because pedestrian injuries tend to cluster in areas with high pedestrian traffic and pedestrian traffic varies widely across New York City, we used a model to estimate a pedestrian count at each audited intersection. Briefly, we calculated estimated counts by obtaining population-density data from the 2010 US Census, commercial zoning data as reported by the New York Department of City Planning, and subway stop and ridership data as reported by the New York State Department of Transportation and the New York City Transit Authority. These data are scaled to approximate a pedestrian count observed over a 10-minute period falling between either 10:00 and 12:00 or 13:30 and 15:30.[25] We combined these counts and spatially estimated them by using kernel density methods to create a raster surface; we summed the average value of the raster surface at blocks adjoining an intersection to estimate pedestrian counts. This approach has been tested for validity against in-person pedestrian counts ($r = 0.62$) within New York City by Purciel et al.[26]

This pedestrian count model contains New York City–specific components such as subway ridership data. To evaluate the information technology approach as it might be used in another city, in contexts in which there are no empirical pedestrian counts to anchor a model, or in a multicity study, we also evaluated the use of Walk Score,[27] a commercially available walkability index shown previously to predict purposive walking[28,29] to estimate pedestrian counts. We had previously obtained Walk Scores at the centroid of each US Census block in New York City; for this analysis, each collision was assigned the nearest available Walk Score. Distances between geocoded collision site and nearest Walk Score ranged from 2 to 238 meters with a mean of 73 meters.

Because pedestrian injury count at an intersection is also affected by the volume of motor vehicles crossing the intersection, many accident models include an estimated average daily traffic count. Unfortunately, no such counts were available for all intersections considered in this analysis. Instead, we used arterial class codes as determined by the New York State Department of Transportation[30] as a proxy for motor vehicle traffic. Specifically, we defined arterials of class 1, 2, 3, or 4 (interstates, expressways, principal arterials, and minor arterials) as major roads, arterials of class 5 or 6 (major collector and minor collector) as minor roads, and classified intersections as minor/minor if no major roads were present, major/major if no minor roads were present, and minor/major if otherwise.

Statistical Analyses

We used the average pair-wise Cohen κ to assess interrater reliability.[31] For ordinal measures (sidewalk and road conditions), we computed weighted kappas by using the square of the difference from perfect agreement as a weight.[32] To assess reliability of pedestrian population estimates, we computed Spearman correlations between modeled pedestrian count and Walk Score values.

Consistent with best practices in traffic safety studies,[33–35] we modeled the count of injury accidents at each intersection as follows:

$$(1)\quad E(Injuries) = \alpha N^{b_0} e^{(\sum b_i x_i)}$$

where N denotes the estimated pedestrian count and x_i ($i = 1,2,3,\ldots, n$) denotes the characteristics of the intersection, including arterial classification codes as a proxy for road traffic, e is the numerical constant base of natural logarithms, and α, b_0, and b_i ($i=1,2,3,\ldots,n$) are estimated parameters. b_0 estimates the relationship between pedestrian volume and count of injuries at a given location such that an estimate less than 1 indicates lower risk per pedestrian at intersections with higher pedestrian volume.[36] Exponentiating the estimated b_i values estimates the percentage difference in count of injuries per unit difference in x_i, holding the pedestrian count constant and conditional on the other built environment characteristics being set to the reference level.

The injury counts were overdispersed (variance-to-mean ratio = 6.08), so we fit the model with negative binomial regression.[37] We first modeled each characteristic separately, then developed a multivariable model that included all characteristics. During multivariable model development, we removed traffic-calming infrastructure from the model because of the rarity of this feature, and we removed the variable for curb cuts owing to their collinearity with pedestrian signals and crosswalks. Likelihood ratio tests

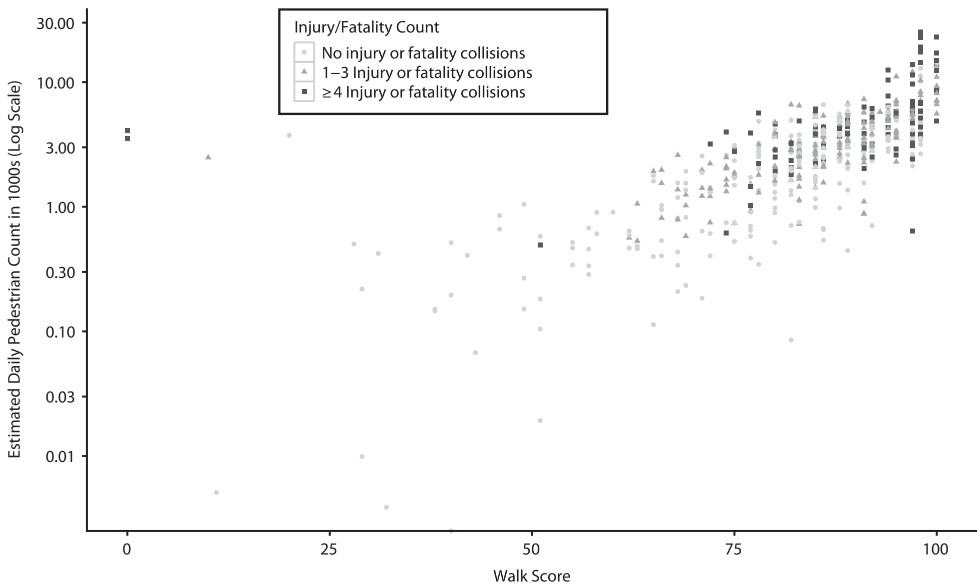


FIGURE 2—Relation Between Log-Transformed Estimated Pedestrian Count and Walk Score for 532 Intersections in New York City in 2013

suggested that all remaining covariates improved model fit (at $P < 0.1$).

New York is divided into 59 community districts whose boards advocate local needs, including traffic safety. As such, injury counts and built environment characteristics at intersections within the same community district may not represent independent observations. For each sampled intersection, we identified the surrounding community district by using a spatial merge. Seven intersections fell outside community district boundaries (e.g., in parks) and were coded as 1 additional district. We computed cluster-robust standard errors to account for this possible nonindependence of observations at intersections within the same community districts.

We performed identification of arterial class codes for streets entering intersections, pedestrian count estimation, and collision count summation in ArcGIS by using the latitudes and longitudes reported by Google Street View at each intersection. We performed all subsequent analyses, including merging with data collected from Street View in R for Windows version 2.15.3 (Vienna, Austria). Interrater reliability analyses used the irr package version 0.84,[38] spatial analyses used the rgdal package version 0.8-14 and sp package version 1.0-14,[39,40] and we generated Figure 1 with the ggmap package version 2.4.[41]

Sensitivity Analyses

We performed 2 sensitivity analyses. First, because the 60-meter buffer we used at intersections might include collisions at nearby intersections in colonial-era neighborhoods in New York City where blocks are short, we performed a sensitivity analysis limited to the 85% of collisions (n = 950) occurring within 30 meters of the intersection center rather than 60 meters.

Second, because it is plausible that interventions (e.g., new crosswalks) observed in 2011 Street View imagery were deployed in response to high injury rates during the beginning of the study period, we analyzed injuries and fatalities that occurred between 2011 and 2013.

RESULTS

Kappa scores for interrater reliability of the items ranged from 1.00 for presence of pedestrian signals to 0.37 for the presence of traffic calming. As expected, Walk Score values were highly correlated with estimated pedestrian counts (Spearman r = 0.76; Figure 2).

Over the 5 years studied, there was a total of 1117 collisions in which a pedestrian was injured (n = 1103) or killed (n = 14) at the 532 audited intersections. The distribution of collision count by intersection was right-skewed. Two-hundred forty-one intersections (45.3% of the overall sample) had no reported pedestrian injury or fatality collisions, whereas the intersection with the maximum injury count had 36. Table 1 displays prevalence of observed pedestrian environment characteristics of all intersections, intersections with any injuries or fatalities, and intersections where fatalities occurred.

Modeling each variable separately resulted in strong, but likely confounded, associations between most environmental characteristics and injury counts (Table B, available as a supplement to the online version of this article at http://www.ajph.org). Only traffic-calming measures, which were rare (n = 11; 2% of all intersections) were not statistically significantly associated with injury count (–32% estimated difference in count; 95% confidence interval [CI] = –74%, 77%).

In the final multivariable negative binomial model incorporating estimated pedestrian counts and using cluster-robust standard errors to account for intersections

TABLE 1—Characteristics of 532 Intersections in New York City Included in This Study as Depicted in Google Street View Imagery: 2007–2011

Characteristic	Estimated Pedestrian Count,[a] Geometric Mean, (Geometric SD)	Kappa Score	Frequency of Characteristic, %	
			All (n = 532)	Any Collision[b] (n = 291)
Crosswalk presence		0.83		
None	1.2 (4.0)		24	12
Connecting some corners	2.3 (2.2)		25	22
Connecting all corners	3.9 (2.1)		48	65
N/A[c]			3	1
Curb cuts		0.48		
Present	2.7 (2.5)		91	96
Not present	1.7 (4.0)		5	3
No sidewalk	0.2 (8.8)		4	1
Visible billboards		0.75		
None	2.3 (3.1)		89	84
≥ 1	3.6 (2.3)		11	16
Sidewalk condition		0.40[d]		
Good	3.2 (2.5)		29	29
Fair	2.7 (2.6)		46	49
Poor	2.0 (2.4)		17	16
Under repair	7.5 (1.9)		2	2
Road condition		0.51[d]		
Poor	3.0 (3.1)		13	12
Fair	2.5 (3.1)		62	62
Good	2.3 (3.2)		24	25
Under repair	3.4 (1.6)		1	1
Pedestrian signal		1.00		
Not present	1.7 (3.4)		55	36
Present	3.9 (2.1)		45	64
Traffic island		0.52		
Not present	2.4 (3.0)		87	80
Present	3.0 (3.1)		13	20
Traffic calming		0.37		
Not present	2.5 (3.0)		98	98
Present	2.5 (4.0)		2	2
Bus stop		0.70		
Not present	2.4 (3.1)		89	84
Present	3.0 (2.2)		11	16

[a]Estimated pedestrian count as observed over a 10-minute period either between 10:00 and noon or 1:30 and 3:30 in the afternoon.
[b]Intersections at which 1 or more injuries or fatalities occurred.
[c]N/A indicates intersections where no crosswalk may be expected, such as at a freeway onramp where pedestrian barriers block access to the roadway.
[d]Weighted κ used for ordinal measures.

nested in community districts, the regression coefficient for the log pedestrian count term was less than 1 (b = 0.53; 95% CI = 0.28, 0.78), indicating that injury incidence per pedestrian was lower at intersections with more pedestrian traffic.

Marked crosswalks (80% increase; 95% CI = 2%, 218%), pedestrian signals (156% increase; 95% CI = 69%, 289%), bus stops (120% increase; 95% CI = 51%, 220%), and billboards (42% increase; 95% CI = 7%, 90%) were associated with increased injury counts. Findings were qualitatively similar between models that used estimated pedestrian count and those that used Walk Score to control for differences in pedestrian volume (Table 2).

Sensitivity analyses limiting buffers around intersection center points to 30 meters and assessing collisions that occurred between 2011 and 2013 resulted in estimates similar to those of the main analysis (Tables C and D, available as supplements to the online version of this article at http://www.ajph.org).

DISCUSSION

This study demonstrates that the use of Google Street View imagery with CANVAS[17,42] and pedestrian count models[26,27] is a viable and informative methodology to identify pedestrian environment characteristics associated with pedestrian injuries. Marked crosswalks, pedestrian signals, visible billboards, and bus stops were all associated with elevated pedestrian injury counts. Consistent with in-person studies,[36,43] injury counts per pedestrian were lower at intersections with higher pedestrian volumes. Models using Walk Score were roughly comparable to those using estimated pedestrian counts. This information technology–based approach obviates the time and expense of in-person visits to crash sites by researchers to collect physical environment data.

Previous studies assessing pedestrian safety have noted that the time required for site visits substantially limits the number of intersections available for study.[44] Using Google Street View to assess intersection characteristics and Walk Score to estimate pedestrian counts may enable a much more efficient data-collection regime with relatively little cost to validity. In one study assessing characteristics of nearly 850 intersections in person,[45] field data collection required slightly more than 3 person-years to complete (T. Koepsell, e-mail communication, December 10, 2014); at 10 minutes per intersection, virtual audit data collection for the same-size study could be completed in slightly less than a person-month.

Though perhaps initially somewhat counterintuitive, our finding that marked crosswalks are associated with elevated injury

TABLE 2—Association Between Intersection Environment Characteristics as Depicted in Google Street View Imagery and the Count of Pedestrian Injuries as Reported to the New York City Department of Transportation at 532 Intersections in New York City From 2007 to 2011, Mutually Adjusted

	Difference in Count of Injuries at Intersection, % (95% CI)[a]	
Characteristic	Model 1[b]	Model 2[c]
Crosswalk presence		
None (Ref)	0	0
Connecting all corners	80 (2, 218)	118 (22, 289)
Connecting some corners	93 (34, 178)	132 (60 ,235)
N/A[d]	−27 (−79, 159)	−11 (−80, 296)
Visible billboards		
None (Ref)	0	0
≥1	42 (7, 90)	54 (18, 110)
Sidewalk condition		
Good (Ref)	0	0
Fair	49 (13, 96)	32 (1, 74)
Poor	53 (8, 116)	35 (−3, 86)
Under repair	44 (−11, 132)	88 (28, 174)
Road condition		
Poor	−38 (−57, −9)	−36 (−56, −8)
Fair	−17 (−35, 6)	−22 (−40, 1)
Good (Ref)	0	0
Under repair	315 (89, 809)	301 (85, 771)
Pedestrian signal		
Not present (Ref)	0	0
Present	156 (69, 289)	182 (81 ,341)
Traffic island		
Not present (Ref)	0	0
Present	31 (−2, 76)	27 (−3, 68)
Bus stop on block		
Not present (Ref)	0	0
Present	120 (51, 220)	103 (37, 201)
Intersection type[e]		
Minor/minor (Ref)	0	0
Minor/major	19 (−11, 59)	19 (−8, 54)
Major/major	3 (−28, 47)	2 (−27, 43)

Note. CI = confidence interval.

[a]Confidence intervals computed by using cluster-robust standard errors, clustering on community district.

[b]Adjusted for log of estimated pedestrian count. Estimated increase in injury count per 1-unit increase in log pedestrian count: 0.53; 95% CI = 0.27, 0.78.

[c]Adjusted for nearby Walk Score. Estimated increase in injury count per 1-unit increase in Walk Score: 0.015; 95% CI = 0.005, 0.025.

[d]N/A indicates intersections where no crosswalk may be expected, such as at a freeway onramp where pedestrian barriers block access to the roadway.

[e]Intersections characterized as minor/minor if no major roads were present, major/major if no minor roads were present, and minor/major otherwise.

counts is concordant with previous findings.[45,46] Although this may be because crosswalks give pedestrians a false sense of security,[47,48] interpreting this association as the intervention effect of adding a crosswalk may not be appropriate. If pedestrians disproportionately choose to cross the most apparently dangerous streets only where crosswalks are present or if the municipal government implements crosswalks at intersections with higher baseline risk, then locations with crosswalks may be associated with elevated injury counts even if the presence of a crosswalk reduces the risk to a given pedestrian who wishes to cross a given street.

Finally, our finding that billboards and bus stops, which frequently include advertising on benches or shelters, are associated with elevated injury counts is consistent with the finding that roadside advertising can contribute to driver distraction.[49] We caution, however, that billboards are typically placed near roads engineered to maximize through traffic, and bus routes usually travel on arterials as well, so billboards and bus stops may also simply mark locations where road engineering is responsible for elevated risk.

Most research on the built environment's contribution to pedestrian risk uses 1 of 2 design types: a "before-and-after" design in which researchers compare the count of collisions before and after interventions at 1 or more sites or a "location-based" design, in which researchers compare pedestrian environments between higher- and lower-risk locations.[10] Because before-and-after designs allow comparisons between the same intersections across time, they control many potentially hard-to-quantify characteristics of intersections by design. However, such studies are subject to regression-to-the-mean artifacts because interventions are often put in place in response to abnormally high numbers of collisions, resulting in overstated impacts from infrastructure improvements.[50] Furthermore, most interventions are deployed gradually to only a small number of intersections at any given time, limiting study to selected locations.

By comparison, location-based designs allow analysis at a broader range of locations and avoid regression-to-the-mean artifacts. However, visits to collision sites to survey environmental characteristics and collect motor vehicle and pedestrian counts can be costly and time consuming to implement.[51] These travel logistics typically limit location-based designs to a relatively few intersections, which may constrain generalizability.[52] By contrast, the "virtual shoe leather" paradigm demonstrated here may be expanded to allow a national scope of inquiry (using sources such as the Fatality Accident Reporting System[53] for collision data) or broader sampling within a local area without incurring logistically cumbersome travel.

We initially assessed several other theoretically relevant intersection characteristics, including traffic signaling and lane counts, but we excluded these from this analysis because data collection artifacts rendered them unlikely to reliably reflect the construct of interest. For example, the protocol used to assess traffic signaling required assessing the signal encountered by traffic approaching an intersection from only 1 direction, whereas injury risk is likely determined by overall signaling. Similarly, lane counts were taken for only 1 street segment entering the intersection of interest. Future work on virtual site visits for pedestrian injury studies will benefit from refinement of audit protocols.

This study's findings are strengthened by the relatively large number of intersections studied and the diversity of pedestrian environments represented across the 5 boroughs of New York City, which range from narrow colonial-era streets in Downtown Manhattan to midcentury suburban-style developments on Staten Island. However, our study has important limitations. First, the temporality of imagery assessed by Google Street View reflects 1 point in time, whereas the collisions accrued over a number of years. Though temporal trends in pedestrian injuries were roughly consistent over the years studied[9] and though assuming an unchanging pedestrian environment is typical of most location-based studies, this disjoint nonetheless introduces potential for measurement error. Indeed, because Google Street View retains images from multiple time points, virtual audits hold promise as a strategy to assess this assumption, potentially in a subsample of a larger study.

Second, like many studies of environmental contribution to collision risks, our analysis treats all pedestrians as equivalent, ignoring substantial between-pedestrian heterogeneity in probability of injury and severity of injury from a given collision.[54] For example, previous analysis has suggested that pedestrians using alcohol are at much higher risk for injury,[55] yet our modeling strategy does not account for differences in probability of alcohol use in pedestrian populations. Future studies may use more advanced statistical techniques to model diverse pedestrian populations or contextual influences such as nearby alcohol outlets. Third, without estimated daily traffic counts at the audited intersections, we were forced to rely on arterial class codes as a proxy for motor vehicle counts, which may have greater error. Fourth, some environment characteristics were not measured reliably, which may have biased findings if errors were correlated with injury risk. Extending CANVAS functionality to flag reliability issues for real-time correction may improve future data collection.

Finally, we selected our study's sample of 532 intersections without reference to street characteristics of interest and as a result it was underpowered to determine if some characteristics of considerable theoretical interest (e.g., traffic-calming measures such as speed bumps and curb extensions) were protective. Future studies may complement the logistical efficiency of virtual location-based studies with more statistically efficient sampling strategies to assess traffic calming features' relation to injury risk more completely.

Although location-based analyses are common in assessment of environmental contribution to pedestrian risk, they cannot be considered to be causal. Case–crossover designs, which remove potential non-comparability of pedestrians, may be more appropriate for assessing the causal effects of intersection characteristics on collision risk. In this design, a trip on which an injury occurred is mapped and characteristics of the location where the injury occurred is compared to 1 or more randomly selected locations where the injury did not occur.[56] Although we did not have individual-level trip data to pilot test such a study, Street View appears well-suited to assess environmental risks in case–crossover studies as well.

In conclusion, use of CANVAS and Google Street View to assess environmental characteristics and exposure assessments such as Walk Score or other models of pedestrian traffic appears to be a promising mechanism to not only reduce costs of, but also increase geographic scope of, location-based studies of pedestrian injury risk. In applying this relatively novel method, we found visual advertising, bus stops, and crosswalk infrastructure to be associated with elevated risk of pedestrian injury in New York City. AJPH

CONTRIBUTORS

S. J. Mooney and A. G. Rundle originated the article, with substantial methodological input from C. J. DiMaggio. D. M. Sheehan assisted with collision data acquisition and performed the spatial analysis. S. J. Mooney, A. G. Rundle, M. D. M. Bader, K. M. Neckerman, and J. O. Teitler designed the environmental assessment system. G. S. Lovasi, M. D. M. Bader, K. M. Neckerman, J. O. Teitler, and D. W. Jack assisted substantially with writing. S. J. Mooney conducted the crash risk analysis and prepared the initial draft. All authors had responsibility for the final content.

ACKNOWLEDGMENTS

Funding for this study was provided in part by the National Cancer Institute (T32-CA09529), the National Center for Injury Prevention and Control of the Centers for Disease Control and Prevention (1 R49 CE002096), the Center for Injury Epidemiology and Prevention at Columbia, and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R21HD062965). S. J. Mooney was partially supported by a fellowship provided by EHE International Inc.

HUMAN PARTICIPANT PROTECTION

The collection of neighborhood data with Google Street View was approved by the Columbia University institutional review board. Locations of pedestrian injuries and fatalities in New York City were taken from a de-identified publically available data set.

REFERENCES

1. National Highway Traffic Safety Administration. Traffic safety facts: 2012 data. 2015. Available at: http://www-nrd.nhtsa.dot.gov/Pubs/811888.pdf. Accessed November 21, 2015.

2. Patek GC, Thoma TG. Commentary: pedestrian fatalities—a problem on the rise. *Ann Emerg Med*. 2013;62(6):613–615.

3. Retting RA. Urban pedestrian safety. *Bull N Y Acad Med*. 1988;64(7):810–815.

4. Beck LF, Dellinger AM, O'Neil ME. Motor vehicle crash injury rates by mode of travel, United States: using exposure-based methods to quantify differences. *Am J Epidemiol*. 2007;166(2):212–218.

5. Wendel-Vos W, Droomers M, Kremers S, Brug J, Van Lenthe F. Potential environmental determinants of physical activity in adults: a systematic review. *Obes Rev*. 2007;8(5):425–440.

6. Carver A, Timperio A, Crawford D. Playing it safe: the influence of neighbourhood safety on children's physical activity—a review. *Health Place*. 2008;14(2):217–227.

7. Muennig PA, Epstein M, Li G, DiMaggio C. The cost-effectiveness of New York City's Safe Routes to School Program. *Am J Public Health*. 2014;104(7):1294–1299.

8. City of New York. Vision Zero Action Plan. 2014. Available at: http://www.nyc.gov/html/visionzero/pdf/nyc-vision-zero-action-plan.pdf. Accessed November 21, 2015.

9. DiMaggio C. Small area spatiotemporal analysis of pedestrian and bicyclist injuries in New York City 2001–2010. *Epidemiology*. 2015;26(2):247–254.

10. DiMaggio C, Li G. Roadway characteristics and pediatric pedestrian injury. *Epidemiol Rev*. 2012;34(1):46–56.

11. Retting RA, Ferguson SA, McCartt AT. A review of evidence-based traffic engineering measures designed to reduce pedestrian–motor vehicle crashes. *Am J Public Health*. 2003;93(9):1456–1463.

12. Tester JM, Rutherford GW, Wald Z, Rutherford MW. A matched case–control study evaluating the effectiveness of speed humps in reducing child pedestrian injuries. *Am J Public Health*. 2004;94(4):646–650.

13. DiMaggio C, Li G. Effectiveness of a safe routes to school program in preventing school-aged pedestrian injury. *Pediatrics*. 2013;131(2):290–296.

14. Griew P, Hillsdon M, Foster C, Coombes E, Jones A, Wilkinson P. Developing and testing a street audit tool using Google Street View to measure environmental supportiveness for physical activity. *Int J Behav Nutr Phys Act*. 2013;10(1):103 .

15. Odgers CL, Caspi A, Bates CJ, Sampson RJ, Moffitt TE. Systematic social observation of children's neighborhoods using Google Street View: a reliable and cost-effective method. *J Child Psychol Psychiatry*. 2012;53(10):1009–1017.

16. Wilson JS, Kelly CM, Schootman M, et al. Assessing the built environment using omnidirectional imagery. *Am J Prev Med*. 2012;42(2):193–199.

17. Rundle AG, Bader MD, Richards CA, Neckerman KM, Teitler JO. Using Google Street View to audit neighborhood environments. *Am J Prev Med*. 2011;40(1):94–100.

18. Badland HM, Opit S, Witten K, Kearns RA, Mavoa S. Can virtual streetscape audits reliably replace physical streetscape audits? *J Urban Health*. 2010;87(6):1007–1016.

19. Mooney SJ, Bader MD, Lovasi GS, Neckerman KM, Teitler JO, Rundle AG. Validity of an ecometric neighborhood physical disorder measure constructed by virtual street audit. *Am J Epidemiol*. 2014;180(6):626–635.

20. Bader MDM, Mooney SJ, Lee Y, et al. Development and deployment of the Computer Assisted Neighborhood Visual Audit System (CANVAS) to measure health-related neighborhood conditions. *Health Place*. 2015;31:163–172.

21. Day K, Boarnet M, Alfonzo M, Forsyth A. The Irvine–Minnesota inventory to measure built environments development. *Am J Prev Med*. 2006;30(2):144–152.

22. Pikora TJ, Bull FCL, Jamrozik K, Knuiman M, Giles-Corti B, Donovan RJ. Developing a reliable audit instrument to measure the physical environment for physical activity. *Am J Prev Med*. 2002;23(3):187–194.

23. Lovasi GS, Bader MD, Quinn J, Neckerman K, Weiss C, Rundle A. Body mass index, safety hazards, and neighborhood attractiveness. *Am J Prev Med*. 2012;43(4):378–384.

24. Lovasi GS, Schwartz-Soicher O, Quinn JW, et al. Neighborhood safety and green space as predictors of obesity among preschool children from low-income families in New York City. *Prev Med*. 2013;57(3):189–193.

25. Neckerman KM, Lovasi GS, Davies S, et al. Disparities in urban neighborhood conditions: evidence from GIS measures and field observation in New York City. *J Public Health Policy*. 2009;30(suppl 1):S264–S285.

26. Purciel M, Neckerman KM, Lovasi GS, et al. Creating and validating GIS measures of urban design for health research. *J Environ Psychol*. 2009;29(4):457–466.

27. Walk Score Inc. About Walk Score. 2014. Available at http://www.walkscore.com/about.shtml. Accessed November 21, 2015.

28. Brown SC, Pantin H, Lombard J, et al. Walk Score: associations with purposive walking in recent Cuban immigrants. *Am J Prev Med*. 2013;45(2):202–206.

29. Hirsch JA, Moore KA, Evenson KR, Rodriguez DA, Roux AVD. Walk Score and Transit Score and walking in the Multi-Ethnic Study of Atherosclerosis. *Am J Prev Med*. 2013;45(2):158–166.

30. New York State Department of Transportation. Functional class maps. 2015. Available at: https://www.dot.ny.gov/divisions/engineering/technical-services/highway-data-services/functional-class-maps. Accessed November 21, 2015.

31. Zwick R. Another look at interrater agreement. *Psychol Bull*. 1988;103(3):374.

32. Cohen J. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychol Bull*. 1968;70(4):213–220.

33. Schepers JP, Kroeze P, Sweers W, Wüst J. Road factors and bicycle–motor vehicle crashes at unsignalized priority intersections. *Accid Anal Prev*. 2011;43(3):853–861.

34. Eenink R, Reurings M, Elvik R, Cardoso J, Wichert S, Stefan C. *Accident Prediction Models and Road Safety Impact Assessment: Recommendations for Using These Tools*. Leidschendam, The Netherlands: Institute for Road Safety Research; 2008.

35. Elvik R. Can a safety-in-numbers effect and a hazard-in-numbers effect co-exist in the same data? *Accid Anal Prev*. 2013;60:57–63.

36. Elvik R, Bjørnskau T. Safety-in-numbers: a systematic review and meta-analysis of evidence. *Saf Sci*. 2015; Epub ahead of print.

37. Miaou S-P. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accid Anal Prev*. 1994;26(4):471–482.

38. Gamer M, Lemon J, Fellows I, Singh P. irr: various coefficients of interrater reliability and agreement (R package version 0.83). 2012. Available at: http://CRAN.R-project.org/package=irr. Accessed November 21, 2015.

39. Bivand R, Keitt T, Rowlingson B. rgdal: bindings for the Geospatial Data Abstraction Library. 2014. Available at: http://CRAN.R-project.org/package=rgdal. Accessed November 21, 2015.

40. Bivand RS, Pebesma EJ, Gomez-Rubio V, Pebesma EJ. *Applied Spatial Data Analysis With R*. New York, NY: Springer; 2008.

41. Kahle D, Wickham H. ggmap: a package for spatial visualization with Google Maps and OpenStreetMap. 2013. Available at: http://journal.r-project.org/archive/2013-1/kahle-wickham.pdf. Accessed November 21, 2015.

42. Anguelov D, Dulong C, Filip D, et al. Google Street View: capturing the world at street level. *Computer*. 2010;43(6):32–38.

43. Elvik R. The non-linearity of risk and the promotion of environmentally sustainable transport. *Accid Anal Prev*. 2009;41(4):849–855.

44. Schneider RJ, Henry T, Mitman MF, Stonehill L, Koehler J. Development and application of the San Francisco Pedestrian Intersection Volume Model. 2013. Available at: http://docs.trb.org/prp/12-4224.pdf. Accessed November 21, 2015.

45. Koepsell T, McCloskey L, Wolf M, et al. Crosswalk markings and the risk of pedestrian–motor vehicle collisions in older pedestrians. *JAMA*. 2002;288(17):2136–2143.

46. Zegeer CV, Stewart JR, Huang H, Lagerwey P. Safety effects of marked versus unmarked crosswalks at uncontrolled locations: analysis of pedestrian crashes in 30 cities. Transportation Research Record. *J Transportation Res Board*. 2001;1773(1):56–68.

47. Mitman MF, Ragland DR. Crosswalk confusion: more evidence why pedestrian and driver knowledge of the vehicle code should not be assumed. Transportation Research Record. *J Transportation Res Board*. 2007;2002(1):55–63.

48. Tidwell JE, Doyle DP. Driver and pedestrian comprehension of pedestrian low and traffic control devices. *Transp Res Rec*. 1995;1502:119–128.

49. Young MS, Mahfoud JM, Stanton NA, Salmon PM, Jenkins DP, Walker GH. Conflicts of interest: the implications of roadside advertising for driver attention. *Transp Res Part F Traffic Psychol Behav*. 2009;12(5):381–388.

50. Elvik R. The importance of confounding in observational before-and-after studies of road safety measures. *Accid Anal Prev*. 2002;34(5):631–635.

51. Miranda-Moreno LF, Morency P, El-Geneidy AM. The link between built environment, pedestrian activity and pedestrian–vehicle collision occurrence at signalized intersections. *Accid Anal Prev*. 2011;43(5):1624–1634.

52. Dumbaugh E, Li W, Joh K. The built environment and the incidence of pedestrian and cyclist crashes. *Urban Design Int*. 2013;18(3):217–228.

53. National Highway Traffic Safety Administration. Fatality Analysis Reporting System. Available at: http://www.nhtsa.gov/FARS. Accessed November 21, 2015.

54. Evans L. Age and fatality risk from similar severity impacts. *J Traffic Med*. 2001;29(1/2):10–19.

55. LaScala EA, Gerber D, Gruenewald PJ. Demographic and environmental correlates of pedestrian injury collisions: a spatial analysis. *Accid Anal Prev*. 2000;32(5):651–658.

56. Teschke K, Harris MA, Reynolds CC, et al. Route infrastructure and the risk of injuries to bicyclists: a case–crossover study. *Am J Public Health*. 2012;102(12):2336–2343.