2013-01-30

Assessing inter-rater agreement of environmental audit data in a matched case-control study on bicycling injuries

Romanow, Nicole T.R.; Couperthwaite, Amy B.; McCormack, Gavin R.; Nettel-Aguirre, Alberto; Rowe, Brian H.; Hagel, Brent E.

BMJ Publishing Group

Romanow, N. T., Couperthwaite, A. B., McCormack, G. R., Nettel-Aguirre, A., Rowe, B. H., & Hagel, B. E. (2013). Assessing inter-rater agreement of environmental audit data in a matched case-control study on bicycling injuries. Injury prevention, 19(5), 336-341.

http://hdl.handle.net/1880/50975

journal article

Assessing Inter-Rater Agreement of Environmental Audit Data in a Matched Case-Control Study on Bicycling Injuries

[*]NTR Romanow, MSc,[+]AB Couperthwaite, BA, [¥]G R McCormack, PhD, [*]A Nettel-Aguirre, PhD, PStat, [+]BH Rowe, MD, MSc, CCFP(EM), FCCP,[*]BE Hagel, PhD

Affiliations:

[*]Departments of Paediatrics and Community Health Sciences, University of Calgary, Calgary, AB, Canada

[+]Department of Emergency Medicine and School of Public Health, University of Alberta, Edmonton, AB, Canada

[¥]Department of Community Health Sciences, University of Calgary, Calgary, AB, Canada

Correspondence to:

Brent Hagel

Departments of Paediatrics and Community Health Sciences, University of Calgary

Alberta Children's Hospital

2888 Shaganappi Trail NW

Calgary, AB T3B 6A8

brent.hagel@albertahealthservices.ca

Tel: 403-955-7403 Fax: 403-955-3055

Keywords: Bicycle, case-control study, psychometric properties

Word count: 2,988

ABSTRACT

Background: Environmental audit tools must be reliable order to accurately estimate the association between built environmental characteristics and bicycling injury risk.

Objective: To examine the inter-rater agreement of a built environment audit tool within a case-control study on the environmental determinants of bicycling injuries.

Methods: Auditor pairs visited locations where bicycling injuries occurred, and independently recorded location characteristics using the Systematic Pedestrian and Cyclist Environmental Scan (SPACES). Two case groups were defined: 1) where a bicyclist was struck by a motor-vehicle (MV); and 2) where the bicyclist's injuries required hospitalization. The two corresponding control groups were 1) where non-MV bicycle-related injuries occurred; and 2) where minor bicycle-related injuries occurred. Inter-rater reliability of each item on the tool was assessed using observed agreement and Kappa (κ) with 95% confidence intervals (CI).

Results: Ninety-seven locations were audited. Inter-observer agreement was generally high (≥95%); most items had a 1-2% difference in responses. Items with ≥5% differences between raters included path condition, slope, and obstructions. For land use, path, and roadway characteristics, κ ranged from 0.3 for presence of offices and cleanliness to 0.9 for schools and number of lanes; overall, 78% of items had at least substantial agreement (κ≥0.61). For bicyclists struck by a MV the proportion of items with substantial agreement was 60%, compared with 73% for non-MV related injuries. For hospitalizations and minor bicycle-related injuries, 76% of items had substantial agreement.

Conclusions: Agreement was substantial for most, but not all SPACES items. The SPACES provides reliable quantitative descriptions of built environmental characteristics at bicycling injury locations.

INTRODUCTION

Research examining the relationship between the built environment and physical activity behavior continues to accumulate.[1, 2] Much of this research relies on instruments that capture different built environment characteristics in order to provide estimates of how conducive environments are for physical activity or walking.[3] These instruments include environmental audits, which involve a systematic observational assessment of characteristics in the physical and social environment (e.g., recreation facilities, sidewalks, roadway characteristics) that support or inhibit physical activity behavior.[4] Audits differ from other tools used in this field (e.g., surveys, interviews, geographical information systems [GIS], desktop mapping) in that audits examine features of a location at the ground-level by direct observation, providing a fine-grained assessment of the environment at a specific location. Researchers, urban and transportation planners, and laypersons can use environmental audit instruments to assess a variety of settings such as neighborhood streets, public open spaces, new developments, and commercial streets. [5] As interest in environmental determinants of physical activity increases, the list of available audit instruments grows. In a 2003 review, 31 audit instruments designed to assess the physical environment for recreation or transportation related walking or cycling were identified.[6]

Despite commonly being used in physical activity research,[3] to our knowledge, no studies have used environmental audit instruments to assess the environmental determinants of pedestrian and bicyclist injury. As such, examining the reliability of these instruments in the injury research setting is warranted. Since many of the existing instruments were developed and tested outside Canada, their reliability and adaptability for use in the Canadian context remains unclear. As part of a case-control study on the environmental determinants of bicycling injuries,[7] we assessed inter-rater agreement of an environmental audit tool.

METHODS

Bicyclist Data collection

As part of a case-control study on risk factors for bicycling injuries, cyclists who presented to one of seven emergency departments (ED) in Calgary (Alberta Children's Hospital, Foothills Medical Centre, Peter Lougheed Centre, Rockyview Hospital) or Edmonton (University of Alberta Hospital, Stollery Children's Hospital, North East Community Health Centre), Alberta were interviewed about their personal characteristics and crash circumstances. These EDs were chosen as they represent all the hospitals in Calgary, and a sufficient number of sites in Edmonton to cover a representative catchment area in that city. The study hospital included designated adult and paediatric regional trauma centres for each city. Anyone who presented to one of the EDs with a bicycle-related injury was eligible to participate.

Patients were identified using the Regional Emergency Department Information System (REDIS), and by screening ED charts daily. Eligible patients were approached by research staff and provided with a study information sheet and consent form. Those willing to participate were interviewed in person in the ED, or on the hospital unit if they were admitted. Eligible patients who were missed in the ED were mailed a study information package and contacted by telephone to request participation. Patients were not eligible if they did not speak English, were cycling indoors, using a stationary exercise bicycle, or were not riding the bicycle (e.g., cleaning the bike) at the time of the injury. All participants provided written or verbal (for telephone interviews) informed consent. Injury information was collected by reviewing participants' medical charts.

Information on crash locations was determined from cyclist responses during the interviews. The case-control status of crash sites was determined based on the type of event, as described by the cyclist, and injury outcomes, as recorded on their medical chart. Two separate case groups were identified. The first case group was composed of sites where a cyclist was injured in a collision with a motor-vehicle (MV). The second case group included cyclists with injuries severe enough to require hospitalization, regardless of what caused the injury to the bicyclist. If a cyclist was hospitalized as a result of a MV collision, they were considered in both case groups. Each case group was examined separately. A separate control group was selected for each case series. To compare with MV cases, the first control group was composed of bicyclists who were not struck by a MV while riding on the road, or not struck by a MV because they were riding on a sidewalk. To compare with the hospitalized cases, the second control group included bicyclists with minor injuries, regardless of riding location (e.g., bike paths) or mechanism of injury (MV or other).

Environmental audits of injury locations

We conducted ground-level environmental audits of 97 locations where injuries were reported to have occurred from May to October 2010. Audits were conducted using the Systematic Pedestrian and Cycling Environment Scan (SPACES) developed by Pikora et al.[8] This instrument has been used in its original form or adapted in several other studies examining the relationships between built environments and physical activity.[9-12]

The SPACES checklist has 37 items (Tables 1 & 2) separated into four sections, each addressing different aspects of the environment (i.e., type of land use, characteristics of the path, roadway characteristics, aesthetics, and accessibility).[8] Data on two sides of a segment (road or path) can be recorded if necessary. The inter- and intra-rater reliability of the SPACES tool was evaluated during the tool development phase, and found to be generally high, even when more subjective elements of the tool were analyzed.[8] The tool requires limited training and administration time, and auditors reported that it was easy to use. No modifications were made to the SPACES before use in our study, and all items listed on the tool were recorded (as applicable). SPACES and the accompanying user manual are publicly available,[13] and consent was given by the original authors to use the tool in the present study.

Twelve research assistants (RAs) participated in a one-day training session (one session in each city) provided by a central research coordinator (NR). The RAs were given copies of the user manual, and were guided through the items included in the audit during classroom training. The RAs practiced using the data collection tool in the field prior to conducting their first audit. Two auditors visited each crash site at the same time of day as the original event, as close as possible to the crash date. The auditors were not always paired with the same person, and as such, the results presented ignore the pairings. One auditor was always blind to the case-controls status of the location to minimize the possibility of introducing recorder bias. Auditors stood at the crash site and independently observed the immediate surrounding area, recording their observations. Auditors agreed on limits for the area to observe, which were determined based on the natural features and design of the location. In general, the location under observation was approximately equal to one street block; however, this varied depending on the type of location. For example, some pathway audit areas may have been shorter due to limited visibility when paths wound around corners.

The types of locations observed included roads, sidewalks, and pathways. Roads were defined as a local, arterial, or divided roads, and highways that accommodate vehicular traffic. Pathways were defined as multi-use (e.g., pedestrian and cyclist) paved surfaces or cyclist only facilities separated from the road. A sidewalk was defined as a pedestrian facility adjacent to a road.

Ethics approval was obtained from the University of Alberta Health Research Ethics Board and the University of Calgary Conjoint Health Research Ethics Board. A letter explaining the project was available for concerned citizens who approached the research team.

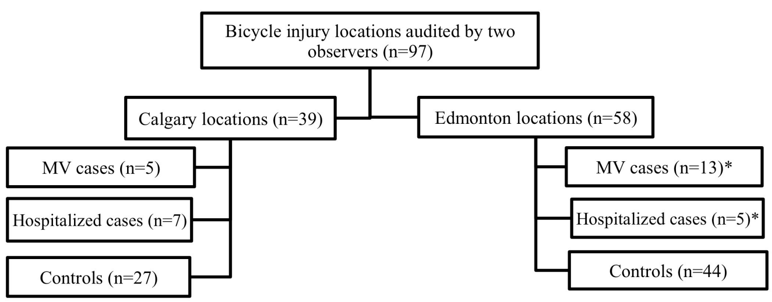
Analysis

We compared the prevalence of each environmental characteristic between auditors. This was calculated as the number of times auditor 1 recorded an item was present, divided by the total number of audit locations. This process was repeated for auditor 2, and the proportions of responses were compared. Kappa (κ) statistics [14] were calculated for individual items on the instrument. Weighted kappa (κ_w) was used for ordinal responses. We also compared κ for cases and controls in order to examine the possibility of observer bias from the non-blinded auditor. Land use and path characteristics were examined separately from roadway characteristics for ease of understanding. The inter-rater agreement estimated using κ was assessed using criteria proposed by Landis and Koch.[15] Based on this criteria, $\kappa < 0.21$ represent poor agreement, 0.21-0.40 fair agreement, 0.41-0.60 moderate agreement, 0.61-0.80 substantial agreement, and >0.80 represented almost perfect agreement. STATA (v.10) statistical software was used for the analyses.

RESULTS

Sample Characteristics

Ninety-seven locations were observed by auditor pairs (Figure 1). Most injury locations were within residential areas (81.1%), where traffic volume was relatively low (49.5% with <250 vehicles/hour). Many incidents involved young cyclists (31.6% less than 13 years old); however, 47.4% involved cyclists 18 years of age or older. There were 4 cyclists struck by a MV and hospitalized in Edmonton. Most events occurred on weekdays (87.4%) between the hours of 16:01 and 20:00 (48.4%), and involved cyclists who had previously cycled at the location of the crash at least 10 times (73.7%). There were 12 (1 hospitalized case and 11 controls) pathway only sites that were not near roads, and the remaining 85 sites were roads, sidewalks, or pathways adjacent to a road.



*Four cases were MV and hospitalised. MV, motor vehicle

Figure 1. Sites Observed for Reliability Assessment in Calgary and Edmonton, Alberta

Observed agreement

Observed agreement was generally high; for most items, there was only a 1-2% difference in the prevalence of specific environmental characteristics. A few items had differences of 5% or more including path condition, slope, and obstructions. The range of differences in responses for these items was from 5.2% for moderate path condition on side 1 to 12.4% for permanent path obstructions on side 2. The presence of street lights and lighting over the path also showed greater than 5% differences in agreement (11.2% for streetlights and 6.2% for lighting over the path). More subjective characteristics or those that required auditors to estimate distances such as tree height, surveillance, and attractiveness, had lower observed agreement compared with more objective items. The difference in the percent of times tree height was recorded as “large” was 9.3%, and the difference for attractiveness was 11.3%.

Agreement measured by Kappa

Observer agreement measured by κ was high for most items. Some of the more subjective observations, such as cleanliness and attractiveness, had fair to moderate levels of agreement. For land use and path characteristics (Table 1), 16 (55.2%) of the 29 items had almost perfect ratings. Eleven (37.9%) were found to have substantial agreement, and 2 (6.9%) had moderate or fair agreement. In Table 2, for roadway characteristics, 25 (65.8%) items had almost perfect or substantial agreement, and 13 (34.2%) had moderate or fair ratings. The range of κ values was from 0.30 (cleanliness) to 0.91 (destinations). Aesthetic characteristics showed the lowest levels of agreement; 2 of the 9 (22.2%) items had almost perfect or substantial agreement, and the remainder had moderate or fair results (77.8%).

Table 1.Kappa Statistics for Land Use and Path Characteristics

	Side 1			Side 2		
	Kappa	95% CI	Agreement	Kappa	95% CI	Agreement
Land use						
Transport	0.83	(0.63-1.0)	Almost perfect	0.82	(0.62-1.0)	Almost perfect
Housing	0.75	(0.55-0.95)	Substantial	0.81	(0.61-1.0)	Almost perfect
Office	0.31	(0.13-0.49)	Fair	0.66	(0.48-0.84)	Substantial
Retail	0.78	(0.58-0.98)	Substantial	0.78	(0.58-0.98)	Substantial
Industry [b]	-	-	-	-	-	-
School	0.92	(0.72-1.0)	Almost perfect	0.92	(0.72-1.0)	Almost perfect
Services	0.83	(0.63-1.0)	Almost perfect	0.85	(0.65-1.0)	Almost perfect
Nature	0.61	(0.41-0.81)	Substantial	0.54	(0.34-0.74)	Moderate
Predominant features	0.79	(0.67-0.91)	Substantial	0.8	(0.68-0.92)	Almost perfect
Same features	0.91	(0.71-1.0)	Almost perfect	n/a	n/a	n/a
Path characteristics						
Type of path	0.79	(0.65-0.93)	Substantial	0.8	(0.66-0.94)	Almost perfect
Path location*	0.88	(0.72-1.0)	Almost perfect	0.88	(0.72-1.0)	Almost perfect
Path material	0.86	(0.74-0.98)	Almost perfect	0.84	(0.72-0.96)	Almost perfect
Path slope*	0.88	(0.72-1.0)	Almost perfect	0.88	(0.7-1.0)	Almost perfect
Path condition*	0.74	(0.6-0.88)	Substantial	0.75	(0.61-0.89)	Substantial
Path obstructions	0.66	(0.54-0.78)	Substantial	0.73	(0.61-0.85)	Substantial

[b]Could not be examined because characteristic was only observed at one location, on one side
*Represents weighted kappa

Table 2. Kappa Statistics for Roadway Characteristics

	Side 1		Side 2		Agreement (side 1, side 2)
	Kappa	95% CI	Kappa	95% CI	
Bike lane	0.65	(0.43-0.87)			Substantial
Road slope*	0.82	(0.64-1.0)			Almost perfect
Road condition*	0.61	(0.45-0.77)			Substantial
Lanes	0.9	(0.76-1.0)			Almost perfect
Parking restrictions #	0.78	(0.58-0.98)	0.70	(0.48-0.92)	Substantial (both)
Curb #	0.85	(0.67-1.0)	0.74	(0.56-0.92)	Almost perfect, substantial
Traffic control	0.85	(0.67-1.0)			Almost perfect
Crossings	0.88	(0.66-1.0)			Almost perfect
Crossing aids	0.7	(0.48-0.92)			Substantial
Other routes	0.6	(0.38-0.82)			Moderate
Street lights #	0.7	(0.5-0.9)	0.65	(0.45-0.85)	Substantial (both)
Lighting covers path #	0.56	(0.4-0.72)	0.53	(0.37-0.69)	Moderate (both)
Destinations	0.91	(0.71-1.0)			Almost perfect
Shop parking	0.85	(0.69-1.0)			Almost perfect
School parking	0.83	(0.67-0.99)			Almost perfect
Other parking	0.8	(0.64-0.96)			Substantial
Bike parking	0.85	(0.69-1.0)			Almost perfect
Driveways*	0.64	(0.48-0.8)			Substantial
Safety characteristics					
Surveillance*	0.57	(0.43-0.71)			Moderate
Garden maintenance*	0.54	(0.38-0.7)			Moderate
Verge maintenance*	0.66	(0.5-0.82)			Substantial
Verge trees* #	0.85	(0.67-1.0)	0.85	(0.67-1.0)	Almost perfect (both)
Tree height* #	0.5	(0.36-0.64)	0.67	(0.53-0.81)	Moderate, substantial
Aesthetic characteristics					
Cleanliness*	0.3	(0.12-0.48)			Fair
Views	0.7	(0.58-0.82)			Substantial
Building similarity*	0.8	(0.62-0.98)			Substantial
Attractive for walking*	0.51	(0.35-0.67)			Moderate
Attractive for	0.48	(0.32-0.64)			Moderate
Difficult for walking*	0.55	(0.35-0.75)			Moderate
Difficult for bicycling*	0.46	(0.3-0.62)			Moderate
Continuity of path	0.48	(0.28-0.68)			Moderate
Neighborhood	0.41	(0.25-0.57)			Moderate

*Represents weighted kappa
#Characteristic was assessed for each side

Reliability for MV case–control groups

We examined reliability by case-control status, primarily to see whether or not there may have been observer bias by knowledge of the status of locations. For MV case and control land use and path characteristics, the proportion of items with at least substantial agreement ratings was higher for controls compared with cases. Twenty-six (89.7%) items had substantial or almost perfect agreement among controls, and 19 (65.5%) did for cases. For roadway items, among controls, 23 (60.5%) of the 38 items had substantial or almost perfect agreement, 13 (34.2%) had moderate ratings, and only 2 (5.3%) were fair. For cases, 21 (56.8%) of the 37 items had at least substantial agreement. When land use, path, and roadway characteristics were examined overall, there was a difference of 13% between the proportions of items with substantial or almost perfect agreement for controls and cases; 73.1% of items for controls, and 60.6% for cases. The items with higher or lower agreement were the same for cases and controls.

Reliability for hospitalized case-control groups

For hospitalized case sites, 19 (70.4%) land use or path items had at least substantial agreement. For control sites where a minor bicycle-related injury that resulted in an ED visit and discharge occurred, 25 (86.2%) items had substantial or almost perfect agreement. For roadway characteristics, the proportion of items with substantial or higher agreement was slightly lower than for path and land use items, but similar for cases and controls. For controls, 26 (68.4%) items had κ values above 0.60. For cases, 27 (73.0%) items had substantial to almost perfect agreement. Overall, 76.1% of items for controls and 76.2% for cases had substantial or higher agreement.

DISCUSSION

Our results confirm the generally high reliability [8] of components of the SPACES tool when used in a case-control study on bicycling injuries in the Canadian setting. In general, observers agreed on the characteristics of locations; observed agreement was above 90% and κ values were in the substantial to almost perfect range for the majority of items. Items that required subjective assessment, such as attractiveness and difficulty for activities, had lower κ values despite a structured training program prior to data collection. In addition, items that required auditors to estimate numbers or distances had lower agreement (e.g., tree height, number of parking spots). In general, roadway characteristics demonstrated lower levels of agreement compared with land use and path characteristics, likely because the former included a greater number of subjective items.

It is useful to compare the results of our reliability testing with those who have previously worked with the SPACES, and also to consider the reliability of other tools based on the SPACES. The results of the initial inter-rater assessment conducted during the tool development phase of the SPACES were similar to our study; lower κ statistics (fair-poor) were found for items in the “views” (e.g., types of views, buildings), “subjective assessment” (e.g., attractiveness, difficulty), and “streetscape” (e.g., tress, maintenance, cleanliness) categories. The authors received formal feedback from the auditors, who reported that it was difficult to assess the attractiveness of an area and to gauge the size and number of trees present.[8] Anecdotally, our auditors reported similar issues. Other audit instruments developed based on the SPACES have had similar reliability results. In general, few items have κ values below 0.40, and as expected, these tend to be questions requiring subjective assessment.[5, 11, 16]

Limitations

Due to limited resources and the availability of auditors, we were able to have two auditors conduct observations at only 97 locations, which represented 35% of the total number of sites observed in the study. This resulted in a relatively small sample for our reliability analysis, particularly when examining subgroups of cases and controls.

For logistical reasons, auditors were not always paired with the same person. The results presented reflect all auditor pairings observed, and therefore, variations in agreement within auditor pairs may be masked. We conducted a separate assessment of the inter-rater agreement for the two pairs that conducted the most audits (results not shown but available from the authors upon request). The results of this assessment suggested that pairs that had more training time with the primary research coordinator classified items more consistently. This is an important point to carry forward in future studies, which should ensure sufficient training time is provided to guarantee auditors are comfortable with the audit instrument.

The reliability analysis by cases and controls showed different results for the struck by MV cases and not struck by MV controls, and hospitalized cases and minor bicycle-related injury controls. A greater proportion of lower agreement ratings for MV cases compared with controls signalled that observer bias may have been introduced by the un-blinded auditor. However, when examining agreement for hospitalized cases and minor bicycle-related injury controls, the overall ratings were similar. The possibility of bias then was only a concern among the MV group. The potential for differential misclassification bias was the primary reason for blinding at least one auditor whenever possible; however, the effect, for example an inflation of the measure of association, of any observer bias would have been minimized by using the blinded audit for further analysis. Another consideration may be that MV case sites had a higher presence of subjective items. We cannot confirm this from the data, but if so, this could in part explain the lower proportion of items with at least substantial agreement among this group. The issue of complexity of locations is something that should be considered in future studies.

This study was conducted in a single Canadian province, and included two similar urban communities. Therefore, the generalizability of the findings may be limited to other areas with similar geography, population density, and urban characteristics. However, these two cities are fairly representative North American urban centres, and as such, the findings could likely be transferred to other similar areas.

While we recognize that the SPACES tool was not initially designed to be used in the context of safety and injury prevention, one of the aims of this study was to examine if it could be. Our results demonstrate that the tool can be successfully and reliably used in this type of research. As such, we do not feel that the original intended use of the toll is a limitation to this study. In fact, the SPACES tool is strengthened by our ability to show that is can be used to collect reliable data in a different research setting than the context in which it was originally developed.

CONCLUSIONS

Our findings suggest that the SPACES tool provides (overall) reliable quantitative descriptions of built environmental characteristics at bicycle injury locations. When items are subjective in nature (e.g., cleanliness), inter-rater agreement diminishes in comparison with more objective items (e.g., path material). As much as possible, audit instruments should focus on using objective measures, or on developing methods of assessing certain features that capture a greater breadth of subjective items (e.g., clarify fixed response categories, provide more fixed-response choices). This information can guide the development of audit instruments for injury research or other applications, and may also inform auditor training approaches, emphasizing the importance of clarifying how to consistently record subjective measures.

CONTRIBUTIONS

NTR Romanow was the primary research coordinator and participated in designing the study, data collection, analysis, and manuscript preparation. AB Couperthwaite was the secondary research coordinator with responsibility for data collection and research assistant support in Edmonton. GR McCormack, A Nettel-Aguirre, BH Rowe, and BE Hagel were part of the study planning and design, and oversaw the analysis phase. All authors assisted in drafting the manuscript and provided critical feedback during manuscript revision.

ACKNOWLEDGEMENTS

The authors would like to thank the clinical and research staff at the Alberta Children's, Foothills Medical Centre, Rockyview, and Peter Lougheed hospitals in Calgary, and the Stollery Children's Hospital, University of Alberta Hospital, and Northeast Community Health Centre in Edmonton. A special acknowledgment is due to the study research staff and students in Edmonton and Calgary, the Paediatric Emergency Research Team (PERT), the volunteers from the Paediatric Emergency Medicine Research Associates Program (PEMRAP) at the Alberta Children's Hospital, and the emergency medicine research assistants at the Foothills Medical Centre.

COMPETING INTERESTS

None declared.

FUNDING

This work was supported by a Grant from the Alberta Centre for Child, Family, and Community Research (09SM-Hagel). Dr. Brent Hagel holds the Alberta Children's Hospital Foundation Professorship in Child Health and Wellness, funded through the support of an anonymous donor and the Canadian National Railway Company, as well as the Alberta Heritage Foundation for Medical Research Population Health Investigator and Canadian Institutes of Health Research New Investigator Awards. Dr. Rowe is supported as a Tier I Canada Research Chair in Evidence-based Emergency Medicine from the Canadian Institutes of Health Research (CIHR) through the Government of Canada (Ottawa, ON). Nicole Romanow was supported by a CIHR Banting and Best MSc Studentship and a studentship from the Alberta Children's Hospital Research Institute for Child and Maternal Health (ACHRI) - CIHR Training Program in Genetics, Child Development and Health Graduate.

REFERENCES

1. Winters M, Davidson G, Kao D, et al. Motivators and deterrents of bicycling: Comparing influences on decisions to ride. *Transportation* 2011;**38**:153-68.

2. Simon DS, Lock F, Lock K. Cycling for transport and public health: A systematic review of the effect of the environment on cycling. *Eur J Public Health* 2010;**21**:738-43.

3. McCormack GR, Shiell A. In search of causality: A systematic review of the relationship between the built environment and physical activity among adults. *Int J Behav Nutr Phys Act* 2011;**8**.

4. Pikora TJ, Giles-Corti B, Bull F, et al. Developing a framework for assessment of the environmental determinants of walking and cycling. *Social Science and Medicine* 2003;**56**:1693-703.

5. Hoehner CM, Ivy A, Brennan Ramirez LK, et al. Active neighborhood checklist: A user-friendly and reliable tool for assessing activity friendliness. *Am J Health Promot* 2007;**21**:534-7.

6. Moudon AV, Lee C. Walking and bicycling: An evaluation of environmental audit instruments. *Am J Health Promot* 2003;**18**:21-37.

7. Ruest N. *Environmental determinants of bicycling injuries*. Calgary: University of Calgary; 2012.

8. Pikora TJ, Bull FCL, Jamrozik K, et al. Developing a reliable audit instrument to measure the physical environment for physical activity. *Am J Prev Med* 2002;**23**:187-94.

9. Evenson KR, Sotres-Alvarez D, Herrin AH, et al. Assessing urban and rural neighbourhood characteristics using audit and GIS data: Derivation and reliability of constructs. *Int J Behav Nutr Phys Act* 2009;**6**:44-59.

10. Bentley R, Jolley D, Kavanagh AM. Local environments as determinants of walking in Melbourne, Australia. *Social Science and Medicine* 2010;**70**:1806-15.

11. Hoedl S, Titze S, Oja P. The bikeability and walkability evaluation table. *Am J Prev Med* 2010;**39**:457-9.

12. Kamphuis CBM, Giskes K, Kavanagh AM, et al. Area variation in recreational cycling in Melbourne: A composition or contextual effect? *Epidemiol Community Health* 2008;**62**:890-8.

13. University of Sydney. *Spaces instrument*. December 22, 2009 [cited 2011 August]; Available from: http://sydney.edu.au/medicine/public-health/cpah/research/spaces.php

14. Viera AJ, Garrett JM. Understanding inter-observer agreement: The kappa statistic. *Family Medicine* 2005;**37**:360-3.

15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159-74.

16. Clifton KJ, Livi Smith AD, Rodriguez D. The development and testing of an audit for the pedestrian environment. *Landsc Urban Plan* 2007;**80**:95-110.