Ph.D. Proposal Draft 3.2

Randall Boyes

August 14, 2018

1 Introduction

Outdoor active play is an important part of childhood development. [1, 2, 3] Outdoor active play reduces sedentary behaviour, protects against obesity, and has an important role in the development of behaviours. [4] Children - both Canadian and international - indicate a strong desire for outdoor active play time, which is characterized by reduced rules and supervision, the ability to interact with other children, and variation in experiences. [5, 6, 7] Canadian children have been engaging in fewer hours of outdoor active play over time. This decrease has been attributed to a number of societal changes, including increasing pressure on parents to protect their children from injury and from abduction, though the latter is quite rare. [8] While there are risks associated with outdoor active play in the short term which are a societal concern, children need physical activity. It is not clear what the risks of outdoor active play are as compared to other forms of physical activity such as organized sport. Further, outdoor active play contributes to child development and this reduction in play may have consequences for health and behaviour as they get older [9].

According to the CDC, the behaviours that pose the greatest risk to adolescent health are substance use, including excessive alcohol use, tobacco use, and drug use; risky sexual behaviours; behaviours that result in injury; physical inactivity; and sedentary behaviour. Outdoor active play reduces sedentary behaviour and protects against obesity. [4] The lessons learned and skills obtained through outdoor active play may also be protective against these other health risk behaviours. By protecting children from injury and the perceived threat of abduction over the short term by limiting their exposure to unsupervised outdoor active play, parents could be exposing children to other long term risks, including chronic disease risk.

Both engagement in outdoor active play in children and health risk behaviour in adolescents are influenced by their physical and social environments. [10] Changes in this environment could be a contributing factor in the observed reduction in overall outdoor active play. Parks, green space, and low-traffic roads provide space to play in, and the social context of the neighbourhood influences both parents' willingness to allow children freedom and children's desire to be outside. But these environments could impact risk behaviour as well. [11, 12] Parks and other outdoor spaces that encourage outdoor active play in children could provide a place for adolescents to engage in risk behaviours. By considering the influence of the environment on both outdoor active play behaviours and risk behaviours, healthy behaviours can be encouraged in all age groups.

1.1 Purpose

This thesis will examine neighbourhood level determinants of outdoor active play and health risk behaviours, as well as relationship between outdoor active play and health risk behaviours at the individual level.

1.2 Objectives & Hypotheses.

The goals of this thesis project are:

1. To investigate the association of a neighbourhood playability index with outdoor active play and health risk behaviour in children. I expect that higher neighbourhood playability will be associated with increases in both outdoor active play and health risk behaviour.

2. To develop an understanding of the injuries that result from outdoor active play relative to the injuries that occur in other types of physical activity. I hypothesize that injuries from outdoor active play will be less severe and less frequent than organized sport injuries.

3. To determine the association of outdoor active play in childhood with adolescent risk taking behaviour. I predict that higher levels of outdoor active play in childhood will be associated with lower health risk behaviour in adolescence.

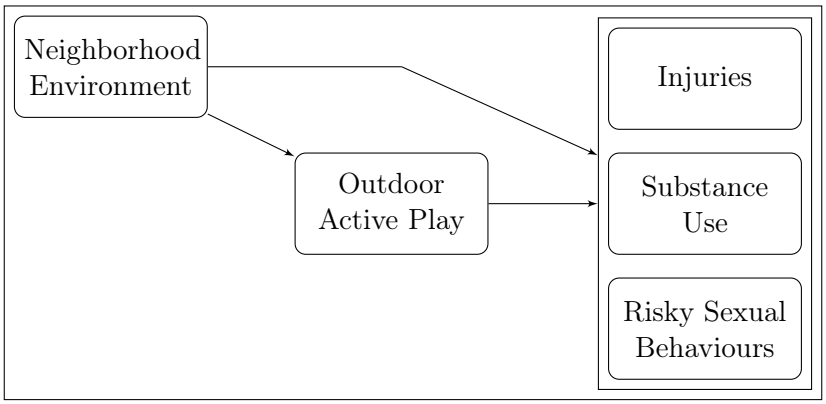


Figure 1: Overall framework. Outdoor active play shapes health risk behaviour (injuries, substance use, risky sexual behaviours) at the individual level. Both outdoor active play and health behaviour are influenced by the neighbourhood environment.

2 Manuscript 1: Influence of the Built Environment

2.1 Introduction

Both physical and social environmental factors influence the behaviour of children and adolescents. [13, 14, 15, 16, 17, 18] Neighborhood social environment in particular has been shown to be one of the strongest predictors of outdoor active play; the social context of the neighbourhood influences both parents' willingness to allow children freedom and children's desire to be outside. [19] The physical environment is also associated with outdoor active play behaviour. Previous attempts to quantify the influence of the built environment on outdoor active play have mostly focused on specific contexts, such as playgrounds, but children engage in outdoor active play in a variety of settings, including - most commonly - near the home. [20]

Increased traffic safety (as measured by the presence of sidewalks (OR 1.44 (1.16, 1.81) for boys, 1.66 (1.39, 1.99) for girls), speed, roundabouts (OR 1.15(1.06, 1.24)), and intersections (OR 0.81 (0.73, 0.90)) in the neighborhood are positively correlated with outdoor active play. [10, 20] Additional neighbourhood characteristics have been shown to be associated with children's physical activity and could relate to rates of outdoor active play as well, including Walkability, number of parks and playgrounds, tight social networks, positive peer influence, and neighborhood safety, recreation facilities, traffic density, walking and cycling infrastructure, both crime-related and subjective safety, and neighborhood aesthetics [21]. +

Neighborhood features have the potential to directly influence health risk behaviours as well.. Hillier hypothesized that there are two classes of space users: 'ordinary citizens', who use space to go about their everyday business; and 'space explorers', whose goals are to explore the potential of space. [11, 12, 22] Children looking for places to play and adolescents looking for places to engage in substance use can both be considered 'space explorers' in Hillier's framework. Therefore, it may be that the very parks and other outdoor spaces that encourage outdoor active play in children could provide a place for adolescents to engage in health risk behaviours.

2.2 Purpose

This initial study will measure the impact of neighbourhood-level features of the built environment as measured by a playability index on both outdoor active play and health risk behaviours.

2.3 Objectives

1. To validate an index of playability that incorporates the features of the built environment of a child's neighbourhood which are associated with outdoor active play.

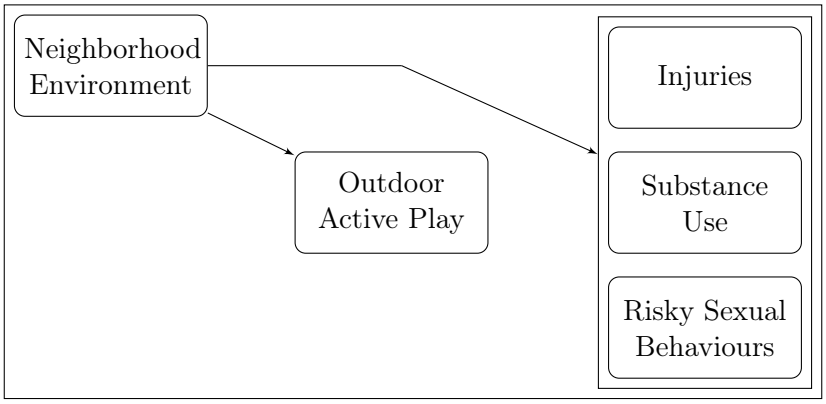


Figure 2: Manuscript one framework. Both outdoor active play and health risk behaviours (injuries, substance use, risky sexual behaviours) are associated with differences in the neighbourhood environment.

2. To determine the association between this playability index and outdoor active play and health risk behaviours after controlling for individual-level confounding variables.

2.4 Study Design and Data Source

Study Design. This is a cross sectional study.

Data Source. The data source is the 2014 cycle of the Health Behaviour in School-aged Children Survey. The 2014 HBSC involved a national survey of 29,784 students from 377 schools from across Canada. A subset of this study (subject to inclusion criteria below) will form the study population.

Inclusion and Exclusion. Inclusions: (1) Participated in the 2014 HBSC; (2) grades 6 to 8 (ages 11-13); (3) attended a school in a public or separate board; (4) reported a valid postal code for their residential address. 16,290 young people met these inclusion criteria.

2.5 Measurement

Exposure: Playability. A neighbourhood playability index, which is being developed by the Playability Project team at UBC, a CIHR-funded interdisciplinary group of researchers led by Mariana Brussoni and Louise Masse, will be tested using the HBSC data. It will include features of neighbourhoods such as number of parks, amount of green space, presence of traffic safety features, and neighbourhood road connectivity. The goal of this index is to define features of neighbourhoods that predict childrens' engagement in outdoor active play. The neighbourhood of a child will be determined using a set distance (1 km) by road from the midpoint of the area represented by their reported home postal code. I will measure the aspects of the playability index that are available in the HBSC data which predict outdoor active play. These features will be measured by primary data collection via geographic information systems analysis, as per precedents in the Janssen laboratory. [23]

Outcome 1: Outdoor Active Play. The number of hours of outdoor active play are assessed on the HBSC using the questions: 'On weekdays, how many hours a day do you usually spend time playing outdoors outside school hours?' and 'On weekends, how many hours a day do you usually spend time playing outdoors outside school hours?' The possible responses to each of these questions were: 'none at all', 'half an hour', '1h', '2h', '3h', '4h', '5h', '6h', and '7h or more.' A total number of hours per week spent in outdoor active play will be estimated from these questions.

Outcome 2: Health Risk Behaviour. The HBSC includes questions on frequency of health risk behaviours, including frequency of binge drinking (5 or more drinks in one day), cigarette smoking, smokeless tobacco use, cannabis use, injuries, and risky sexual behaviours. Tests of self reported measures like the ones used here for substance use versus urinalysis show sensitivity and specificity of approximately 80%. [24, 25, 26] Self reported smoking measures

have even higher sensitivities and specificities of approximately 90%. [27] These measures of these behaviours have previously been combined to form a single continuous score. [28]

Confounders and Effect Modifiers: Sex, age, certain disabilities, rurality, ethnicity, self-reported SES, sleep, family structure, and family attachment can be measured using the HBSC. See Section 5 for evidence of relationships and below for selection process.

2.6 Statistical Analysis:

2.6.1 Objective 1: Validation of the Playability Index.

Qualitative results from the Playability project will be tested using HBSC data. The purpose of the Playability Index is to correlate with the rate of outdoor active play for any given neighbourhood. In order to avoid overfitting, model fit will be checked in bootstrapped samples. Different modelling strategies will be tested to find the best strategy. Models will differ by variable inclusion and functional form of the included variables. As we are aiming to generate a continuous score with an interpretable model, possible models include penalized GLMs such as lasso or ridge regression.

2.6.2 Objective 2: Associations with the Playability Index.

Assumptions for Causal Inference: Making causal inferences about neighbourhood factors in this context is difficult; in fact, arguments have been made that suggest that causal inference in this context cannot be done. [29] Such arguments rest on (amoung other things) violations of the exchangeability and positivity assumptions - that is, it is believed that "a similar person in a different neighbourhood" does not exist due to self-selection of participants into neighbourhoods. Causal inference may still be possible due to the population of interest being children, who typically do not choose their neighbourhood.

Modelling Strategy: As the HBSC is uses a clustered sampling approach, the models of the associations between neighbourhood playability and outdoor active play and health risk behaviours will need to be mixed effects models. Random effects for neighbourhood, school, and class will be used to correct the estimates of variance. Effect estimates will be relative risks for binary outcomes (using log-binomial mixed effects models) or slopes for continuous outcomes.

Power: Assuming a sample size of 16,290 spread over 377 clusters with estimate of 0.10 for the intraclass correlation coefficient based on previously calculated intraclass correlation coefficients for binge drinking and smoking in this data set for 15 year olds [30], the design effect can be calculated as:

$$D = 1 + \rho(k - 1)$$

$$D = 1 + 0.10(16290/377 - 1)$$

$$D = 5.22$$

If a simple comparison was done between high playability and low playability neighborhoods, the minimum detectable effect would be 0.91 standard deviations at 80% power, 1.06 standard deviations at 90% power, or 1.17 standard deviations at 95% power. [31] For categorical variables, 80% power is attained at a true relative risk of approximately 1.35.

Covariates. In order to estimate the relationship between neighborhood factors and outdoor active play behaviour, the models will need to control for both neighborhood- and individual-level confounding variables. Confounder selection can be based on the minimal subset of variables that close all back door paths from exposure to outcome as long as an accurate DAG can be drawn. [32] Plausible causes of the built environment in which children grow up include physical health, socioeconomic status, ethnicity, and family structure, attachment, and environment. These variables could be causally related to neighbourhood built environment by influencing the parent's choice of where to live. Variables such as sleep duration, sex, and age, which are not causes of built environment factors, will not be controlled for. Effect modification by sex and age will be investigated. More detail and some additional potential covariates are listed in Section 5: Potential Covariates.

Sensitivity Analysis by Rurality. Because exposure information is being collected by postal code, assessment of the neighbourhood where a child lives will be more accurate in urban areas than in rural areas. A sensitivity analysis will be conducted where the sample is restricted to non-rural areas, as defined by Statistics Canada. [33]

3 Manuscript 2: Outdoor Active Play and Injury

3.1 Background and Rationale

Caregivers limit childrens' freedom to participate in outdoor active play because they perceive this behaviour as being unacceptably dangerous, though the degree of concern depends on child-level factors including age, competence, and gender. [8, 34] However, there is a lack of basic descriptive data on injuries that occur during outdoor active play. By quantifying this risk and comparing it to other common activities, we can enable decision making at both the social and parental level.

Rates of injury in the context of the playground has been studied extensively, but there is little agreement between studies. A systematic review of play injuries in child care centres found injury rate estimates varied from 0.006 to 8.29 injuries per child-year, depending on the definition of "injury." [35] Outdoor injuries in general in children have been shown to result from play, sport, and transport, with trampolines, bicycles, and skis being frequent causes. [36, 37, 38]

3.2 Purpose

This study aims to describe outdoor active play injuries and to compare the patterns of injury in outdoor active play with those experienced during alternative forms of physical activity.

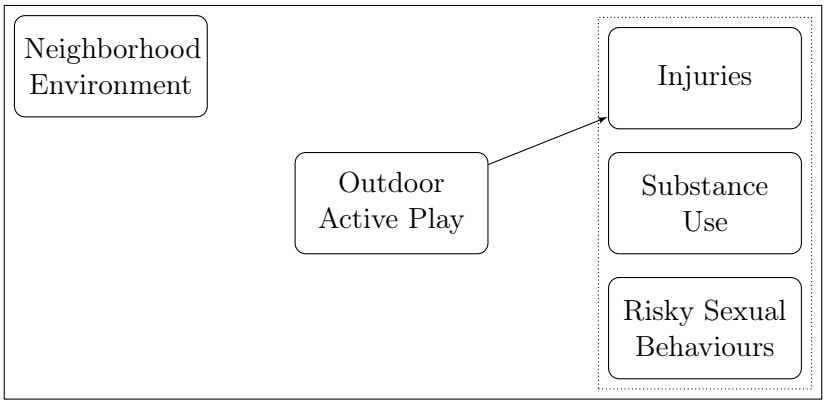


Figure 3: Manuscript two framework. Outdoor active play can lead to directly to injuries in the short term.

3.3 Research Question(s)

What types of injuries occur when children engage in outdoor active play? Are these injuries different in severity, frequency, anatomical site, or nature from the injuries that occur when children engage in more structured activities (i.e., active transport, school-based organized activities, or organized sports)?

3.4 Study Design

Descriptive.

3.5 Inclusion and Exclusion.

Children aged 5 to 14 who presented to hospital with an injury resulting from physical activity in Kingston between 2008 and 2018. This age range was chosen for developmental relevance and to align with census age categories. [39, 40].

3.6 Data Source

Data describing injuries that presented for medical care in Kingston and the surrounding area will be obtained from the Canadian Hospital Injury Reporting and Prevention Program (CHIRPP) database, which captures injuries in patients reporting to emergency departments of the Kingston Health Sciences Centres (formerly Hotel Dieu and Kingston General Hospital). CHIRPP includes information on: 1. What was the injured person doing when the injury occurred? 2. What went wrong? 3. Where did the injury occur? and 4. The nature and anatomical site of the injury.

Injury Rates. In order to calculate rates, both the numerator (the number of injuries resulting from each type of physical activity) and the denominator (the number of children of each age and sex who were at risk) need to be measured.

Number of Injuries. Injuries that result from physical activity, including outdoor active play, will be identified based on open and close-ended codes describing where the injury occurred and what the child was doing, as well as narrative case descriptions. Participants are asked to indicate whether the injury in question happened indoors or outdoors and to indicate the specific activity they were participating in from a list which includes "playing" as an option. We expect that approximately 80% of outdoor active play-related injuries will be identified using these questions, however, cases will exist that are coded as "unknown." For these cases, identification will be attempted using the narrative entry (Section 5). Automated methods exist for topic detection in this type of narrative, such as Latent Dirichlet Allocation[41], and will be used if necessary. Latent Dirichlet Allocation has been used in the past to extract topics from patient narratives [42, 43] Simpler methods involving keywords can also be used. [44] All cases will be classified based on the description of where the injury took place, what the child was doing, and who they were with, as available. The injury classification system will need to be developed as part of this thesis project, which represents a substantial time commitment.

Number of Children at Risk. This study will focus on data from the Kingston Health Sciences Centre. [45] The catchment area of these two hospitals is approximated by the census population of the city of Kingston, Frontenac county, and Lennox & Addington county. [45] Population by age and geographic region cross-tabulations are publicly available for the regions in question from the 2006, 2011, and 2016 censuses. [46, 40] Populations at the years between censuses can be approximated via linear interpolation. This allows the number of person-years at risk in each age/sex strata to be determined. Unfortunately, we cannot determine risk per hour of activity, which means that comparisons will be done with cumulative incidence rates rather than rates per person-hour.

3.7 Presentation and Comparisons of Interest.

Estimates. Injuries will be described according to nature and anatomical site of injury; activity leading to injury; location where the injury occured; and severity of injury. Descriptive tables will be created to show frequency of injury by category of physical activity. Patterns by age, sex, and location (if available) will also be reported. Illustration of typical or recurrent patterns of injury with vignettes will be explored as appropriate. [45]

Precision. If there are 2000 injuries per year, split evenly between injuries in outdoor active play, sport, active transport, and school based activity, and between sexes, then we would observe 250 events in each sex/type stratum per 16,285 person years (2016 Kingston population ages 5 - 14). This would correspond to a incidence rate estimate of 1.54 (1.35, 1.73) per 100 children per year (poisson confidence interval). The rates of injuries are not expected to all be equal. If there is a category with half the expected rate, its rate estimate would be 0.76 (0.63, 0.90) per 100 children per year.

4 Manuscript 3: Outdoor Active Play and Future Health Behaviour

4.1 Rationale

Outdoor active play is an important part of children's development. Outdoor active play helps to form ideas, teach skills, and shape behaviour; some of which may carry forward into adolescence. Outdoor active play may have different

effects on behaviour at different ages. Studies in neuroplasticity suggest that the brain is most susceptible to changes when we are young. [47]

According to the CDC, there are six main categories of behaviour that endanger health in adolescence: 1. Behaviours that contribute to unintentional injury and violence, 2. sexual behaviours that lead to unwanted pregnancies and STIs, 3. use of alcohol and other drugs, 4. tobacco use, 5. unhealthy dietary behaviours, and 6. inadequate physical activity. The first four of these behaviours have been observed to cluster together, suggesting that there may be common underlying causes for all of them. [28, 48]

Problem Behavior Theory [49] suggests that different overt risk taking behaviours (substance use, violence, and risky sexual behavior) are motivated by similar personality characteristics and social influences. [50] This theory is supported by the fact that these overt risk behaviours have been shown to correlate with each other in the HBSC. [28] While it is not known for certain which specific personality characteristics account for this connection, there are some candidate characteristics that are also influenced by outdoor active play. These factors include incorrect evaluations of risk [51], an external locus of control, decreased executive function, an inability to connect with peers in a healthy way, and poor strategies for coping with stress. [52, 53, 9]

4.2 Purpose

This study will longitudinally examine the effect of childhood play on adolescent risk taking behaviours.

4.3 Research Question(s)

1. Is outdoor active play in childhood influence associated with health risk behaviours in adolescence? If so, at what age is this association strongest?

2. Has this association between outdoor active play and health risk behaviours changed over time?

4.4 Study Design

Secondary Analysis of a Prospective Cohort Study.

4.5 Data Source

This study will be a secondary analysis of data from the National Longitudinal Survey of Children and Youth (NLSCY), a 14 year longitudinal study of young Canadians. Cycle 1 data were collected in 1994 and consisted of 16,903 children aged 0 - 11 years. This group was surveyed again every two years until 2008 (remaining n = 10,268). This is the most recent longitudinal Canadian data source for this age group that includes the questions needed, but children today may not be the same as the children who participated in this study.

4.6 Inclusion and Exclusions

All children who responded to the original survey and remained in the cohort for at least one outcome observation will be included in this study. The population of the original survey includes non-institutionalized civilians residing in one of the ten Canadian provinces. The survey excludes children living on First Nations reserves or Crown lands, residents of institutions, children of full-time members of the Canadian Armed Forces, and residents of some remote regions. [54]

4.7 Measurement

Exposure: Outdoor Active Play. If the children were between 4 and 9 years of age, their PMK (Person most knowledgeable) was asked: 'In the last 12 months, outside of school hours, how often has [child's name] taken part in unorganized sports or physical activities?' This was changed to '...taken part in unorganized sports or physical activities *without a coach or instructor*?' and the age was modified to only include 6 to 9 year olds in Cycle 2. In Cycle 4, the minimum age was reduced to 3, where it remained until the final Cycle in 2008. Additionally, children aged 10 and older were asked to indicate how often: 'Outside of school, I play sports or do physical activities WITHOUT a coach or instructor.' The wording was updated in Cycle 2 to include '(biking, skateboarding, etc.)' Children ages

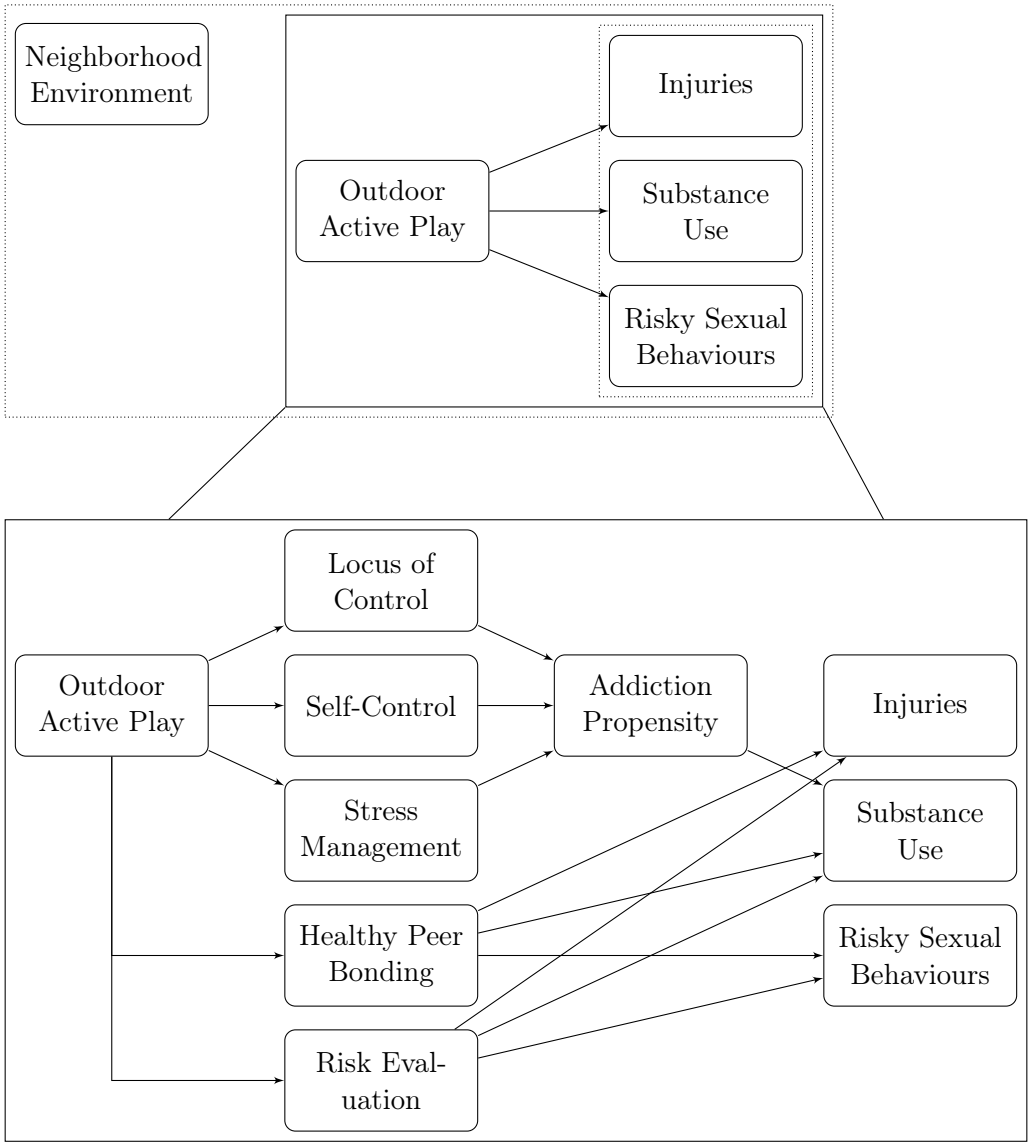
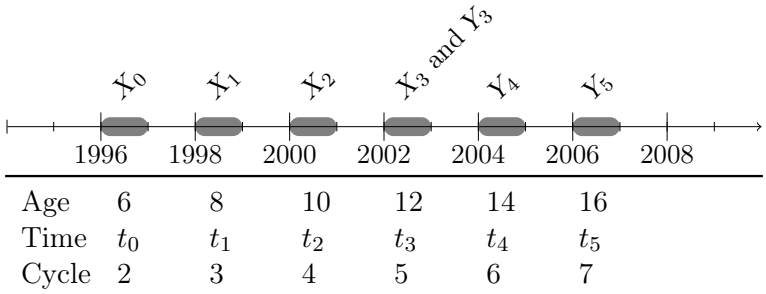


Figure 4: Potential mechanisms behind of the relationship between outdoor active play in childhood and risk taking behavior in adolescence. Outdoor active play could influence childrens' locus of control, self-control, stress management and/or peer bonding strategies, or evaluation of risk. These traits influence attitudes towards health risk behaviours.

14 and older are not asked about play outside of school. The typical child in the dataset will have their frequency of outdoor active play measured at 4 time points with 2 years between measurements (See Figure 5 for example).

Outcome: Health Risk Behaviour. For cycles completed between the ages of 13 and 17, the participants filled out a questionnaire that asked questions about health risk behaviour including substance misuse, failure to take health precautions, and engagement in activities that place the child at physical or emotional risk (e.g. illegal activities). The questions on substance use behaviour come from the Youth Smoking Survey, the HBSC, the Northwest Territories Health Attitudes, Knowledge, and Behaviours Study, the Western Australia Child Health Survey, and the National Longitudinal Survey of Youth at Ohio State University. Tests of self reported measures of substance use like the ones used here versus urinalysis show sensitivity and specificity of approximately 80%. [24, 25, 26] An overall health risk behaviour score similar to the one derived from the HBSC questions in Manuscript 1 will be derived from response patterns on these questions. [55, 56]

Potential Covariates: See Section 5 for potential confounders and effect modifiers. These measurements are based on self-reported data from the NLSCY questionnaire.



Age	6	8	10	12	14	16
Time	t_0	t_1	t_2	t_3	t_4	t_5
Cycle	2	3	4	5	6	7

Figure 5: Exposure and outcome measurement for a single example individual.

4.8 Statistical Analysis

Care will be taken to ensure that the results obtained from models are interpretable. High priority will be placed on clear explanations of findings through figures and simple tables. Confounder selection will be based on an attempt to select the smallest subset of variables that is sufficient to close all backdoor paths from exposure to outcome. The effects of misclassification and missing data will be considered.

4.8.1 Objective 1: Is play related to Health Risk Behaviour?

Sequential conditional mean models [57] will be used to estimate the total effect of outdoor active play on health risk behaviours. A separate model will need to be fit for each combination of outdoor active play at each measured time point (X_0 to X_3) and the outcome at each time point (Y_3 to Y_5), controlling for confounders at the same time as the exposure of interest (L_0 to L_3). This approach relies on the assumption that outdoor active play after t_3 either does not occur or has no effect on the outcome. Generalized estimating equations can be used to estimate the coefficients, but prior exposure and outcome history need to be controlled for to obtain unbiased estimates. [57] The models are then of the general form:

$$E(Y_t|X_{t-a}, L_{t-a}) = \beta_0 + \beta_{X_2} X_{t-a} + \beta_{X_1} X_{t-a-1} + \beta_Y Y_{t-1} + \beta_{L_{t-a}}^T L_{t-a}$$

For t = (3, 4, 5) and a = (1, 2, 3, 4), except for the combination t = 3 and a = 4. These 11 models represent the total effect of outdoor active play at each time point on health risk behaviours at each time point. In addition, if a propensity score can be estimated which balances covariates, then the models will control for this score as well. A propensity score would need to model the likelihood of engaging in outdoor active play, and would include variables such as age, sex, ethnicity, and SES. Controlling for propensity score gives doubly robust estimators of the exposure effect and down-weights exposed individuals who have no comparable unexposed counterparts. [57] Estimates of effect could be presented in a figure similar to Figure 6.

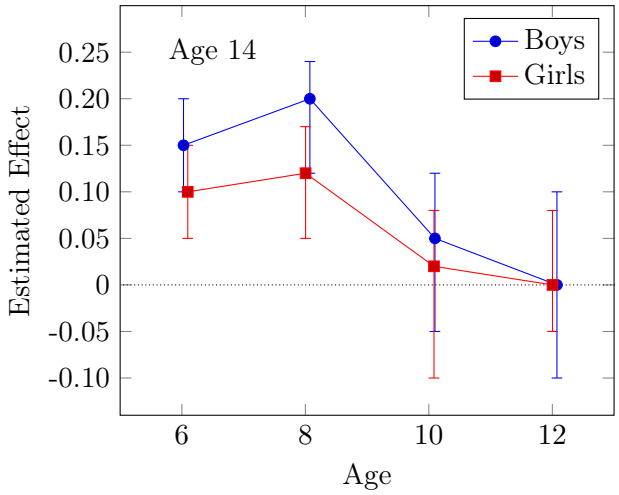


Figure 6: Example plot of associations between exposure at all times and outcome at a specific time point, stratified by sex.

4.8.2 Objective 2: Age, Period, and Cohort Effects

There are three potential concurrent temporal effects in the data. 1. The age of the child is increasing. Differences in effect by age are what we wish to measure. 2. The year of measurement varies. If changes in the built and social environment external to the child interact with the effect of outdoor active play, then we would expect the year of measurement to be important. Such changes can affect all measurements taken in a given year, regardless of age. This type of effect is referred to as a 'period' effect. 3. People born in different years may be fundamentally different. This is a 'cohort' effect. These three effects suffer from the 'identification problem.' That is, fundamentally:

$$\text{Age} = \text{Period} - \text{Cohort}$$

And so there is logically no statistical procedure that can separate their effects without further assumptions being made. Graphically, this can be seen in Figure 7: 'Ten year olds who were measured in Cycle 1' and 'Ten year olds who were born in 1984' are equivalent groups. One potential assumption that can be used is to assume that long term period effects have no trend, which effectively assumes that any trends that do show up in the data are cohort effects. [58] This assumption is not without costs; intervention strategies differ based on whether temporal trends are assumed to be cohort effects (the only effective interventions occur early in life) or period effects (interventions could be effective at any point). Given this assumption, the analysis described for objective 1 can be repeated in subgroups by birth year to examine changes over time. Any changes would be interpreted as cohort effects.

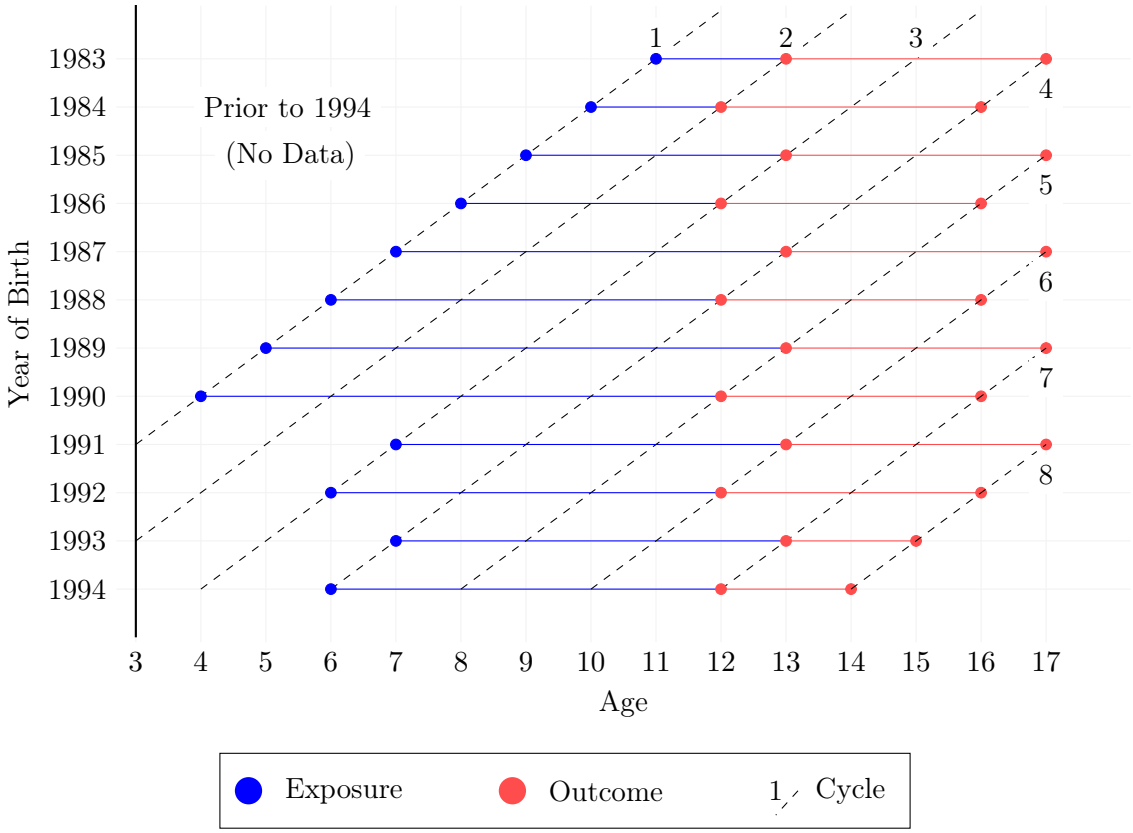


Figure 7: Data availability for exposure and outcome depends on year of birth (which determines age at entry to the cohort). Diagonal dashed lines indicate cycles of the survey. Blue represents time when play can be measured. Orange represents time when health risk behaviours can be measured. Measurement occurs at the intersection of the dashed lines and the coloured lines.

4.8.3 Missing Data

Three separate mechanisms for missing data exist in this dataset.

Systematic Missing Data by Birth Year. As shown in Figure 7, there is a systematic lack of data on exposure for children with earlier birth years and a lack of some outcome data for children with later birth years because of how the survey was conducted. A sensitivity analysis could be conducted by restricting the sample to children born between 1987 and 1992.

Loss to Followup. Over the course of the study, approximately one third of the participants were lost to followup. If the participants who drop out of the study have a different association between health risk behaviour and outdoor active play than those who remain, estimates of effect could suffer from a selection bias. Comparisons of baseline characteristics of these two groups could give an initial idea of the extent of this bias.

Question Non-Response. As with any survey, some individual questions will not be answered by participants. Multiple Imputation can be used to recover these observations as long as the data are missing at random (conditional on observed covariates). Questions on health risk behaviour could be missing not at random (MNAR) if participants who (for example) have tried smoking differentially refuse to answer the question.

4.8.4 Assumptions for Causal Inference

In order to estimate the causal effect of childhood outdoor active play on adolescent health risk behavior, we need some conditions to be true. [59] The assumptions and some comments on their plausibility follow:

Exchangeability requires that children with low outdoor active play and high outdoor active play are otherwise exchangeable (conditional on some set of confounders L). This condition means that - for all two-group sets of children with identical levels of all variables L - switching their mean exposure would switch their mean outcome. Unmeasured confounding and selection bias are threats to exchangeability. In this study, there are some possible confounders for which we do not have a measure: childhood maltreatment and residential mobility. Analysis without these covariates will require that we assume they have a small effect on exposure or outcome, that the are very uncommon, or that they do not enable an unblocked backdoor path from exposure to outcome.

Positivity requires that the distribution of covariates L in children with low outdoor active play completely overlaps the distribution of L in children with high outdoor active play. Testing of this assumption can be accomplished via parametric bootstrapping and deviations can be addressed by reducing the number of parameters in the model, redefining parameters in the model, or restriction of the sample. [60]

Well Defined Intervention. The numerical value of the effect estimate depends on what are considered to be the intervention and control conditions. This is especially important in longitudinal data. The proposed causal effects of interest are the total effect of play at specific time points. Total effects include the direct effect of the exposure at the time point in question as well as the effect that that exposure has though future exposures which are caused by it; in this case, if play at time 1 has an effect on play at time 2, then some portion of the effect of play at time 2 on health risk behaviours should be attributed to play at time 1. This formulation allows estimation of the overall effect of a hypothetical intervention which increased play at any given time point.

Measurement Error and Model Misspecification. Measurement error in the treatment, outcome, or confounders can result in bias in the causal effect estimate. The most likely source of measurement error in this study is on the exposure, which is measured once for each two year period and then extrapolated to the remaining weeks. Measurement of the confounding variables will be imperfect as well. Multiple models will be fit and compared, but perfect model fit can never be guaranteed.

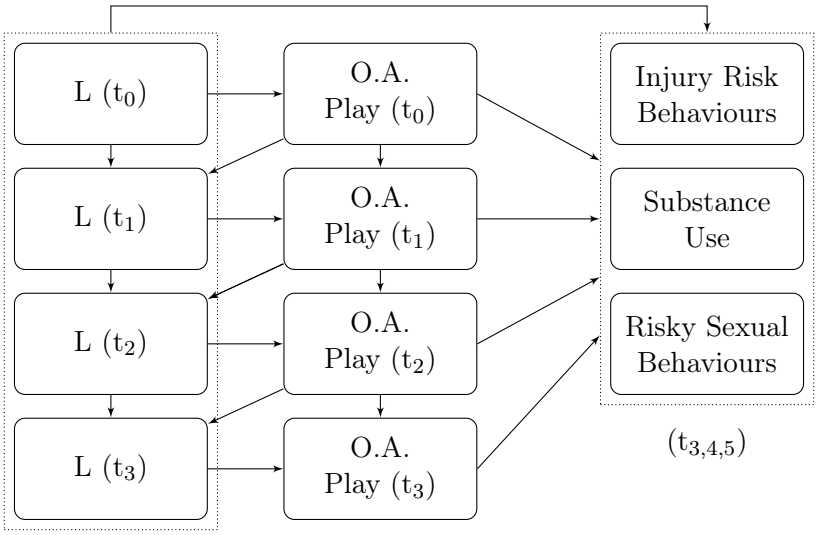


Figure 8: Relationship of play and the set of confounders (L) from t_0 (Age 6) to outcome assessment ($t_{3,4,5}$). Individual's confounder set L changes over time as a result of both play and external factors. This confounder set influences both future play and eventual outcome.

4.8.5 Time-Varying Confounders

Ethnicity and sex are constant, but all other variables are potentially time-varying. Changes in these variables could be related to outdoor active play activity, either as a cause or an effect (See Figure 8). Models will incorporate changes in these covariates over time if necessary.

4.8.6 Sampling Variance

The NLSCY includes a standard set of bootstrap weights for use in calculating variance inflation due to the survey design. Procedures outlined in the NLSCY microdata user's guide will be followed [61].

4.8.7 Power and Sample Size

Simulations indicate that with 10,268 sets of 4 observations and a continuous outcome, 80% power would be reached if the difference between a overall high play and overall low play was greater than 0.33 standard deviations on the outcome scale (assuming equal standard deviation in both groups).

5 Potential Covariates.

Confounding variables will be selected based on past literature support of association with exposure or outcome and availability in the dataset. A preliminary list of variables that may be associated with outdoor active play and/or health risk behaviours follows:

5.1 Effect Modifiers

Sex. Sex will be considered as an effect modifier because boys and girls have both different play experiences and different motivations for health risk behaviour. [62] Boys and girls also interact with their environment in different ways. [63, 64, 65, 66] Boys spend more time in general in outdoor active play [67, 68] and are more prone to health risk behaviors than girls. [69, 70] Boys are significantly more likely to engage in multiple risk behaviors than girls. [71]

Age. Increasing age is linked to decreasing outdoor active play and physical activity [67, 72]. The long term behavioural effect of outdoor active play is also expected to fall off as children age as younger children have higher neuroplasticity. [73] Younger people are more likely to display multiple risk behaviors,[71] but substance use is more common in older adolescents. [70]

5.2 Measured Confounders

Physical Health. Physical disabilities are associated with reduced outdoor active play [74] Chronic pain disorders associated with substance use. [75] Children with physical health issues interact with their environment differently. [76, 18, 77] The effect of ethnicity, income, education on substance use depends on disability status in adults. [78]

Rurality. Rurality is associated with an increase in outdoor active play of .28 hours/day. [68] Rural adolescents have a 25.8% higher rate of binge drinking than urban adolescents and may initiate substance use earlier. [79] There are also small but non-significant differences in the rates of other substance use. Decision making around substance use in rural males depends on group membership and perceptions of health risks - with users and non-users disagreeing on the immediacy, severity, and inevitability of the health risks of use. [51] The suitability of neighborhoods for play varies between urban and rural neighborhoods; for example, traffic speeds tend to be higher in rural neighbourhoods. [80]

Ethnicity. Ethnicity influences attitudes toward violence [81] and is associated with physical inactivity and sedentary behaviour[72, 82] and substance use, with odds ratios vs. Caucasian rangin from 0.55 - 1.58 for binge drinking and from 0.73 to 1.63 for smoking. [83] It is also correlated with outdoor active play behaviours. Amounts of time spent in outdoor active play vary by approximately 0.5 hours per day between the highest and lowest groups. [68] Access to neighbourhood play spaces varies by ethnicity in the United States, with hispanic children having the most and white children having the least [84], but it is unclear whether this pattern would generalize to a Canadian context.

Socioeconomic Status. Children from medium socioeconomic status households engage in more outdoor active play than those from low or high socioeconomic status households (+15.6 min/day (0.3, 30.9))[67]. Lower socioeconomic status is linked to alcohol use (OR 2.63(1.39, 4.95))[85], smoking [70], marijuana use (OR 2.16 (1.32, 3.53))[85], and multiple risk behavior (2.63 (2.12, 3.26))[71]. Attempts to quantify the quality of outdoor active play spaces in low and high SES neighborhoods show dramatic differences on almost all measures of quality. [86] However, areas of with low SES don't have significant differences in parental perception of risk. [87]

Sleep. Short sleepers have reduced hours of outdoor active play. [88] Insufficient sleep is also associated with increased smoking, alcohol use, and drug use in rural adolescents in the US. [89] The physical environment impacts sleep through traffic noise (High traffic is associated with a decrease of 19.2 min/night of sleep vs. low traffic) or through excessive light at night. [84]

Family. Children in single-parent households play 9.9 min/day (2.2, 22.4) more than children in multi-parent households. [67] Number of siblings is not associated with increased outdoor active play time. [90] Higher maternal stress is associated with reduced outdoor active play. [91] Adolescents in secure parent-adolescent dyads are less likely to engage in risky sexual behaviours and substance misuse. [92] Parallel process latent growth models indicate parenting style influences externalizing behavior including violence and substance misuse [93]. Having brothers can increase the likelihood of substance use. [94] Adolescents with low family attachment have higher substance use and higher rates of risky sexual behavior. [95, 94] High parental involvement reduces rates of substance use. [94] Parents' perception of neighborhood safety increases outdoor active play time. [96]

Mental Health. Poor mental health is known to reduce childhood physical activity and could reduce outdoor active play as well. [97] Poor mental health is associated with increases in risky sexual behaviour and substance use. [75] Poor mental health could be influenced by the built environment. [98, 99]

5.3 Unmeasured Confounders

Residential Mobility. Children whose parents move more often are more likely to engage in overt risk behaviours. [100] However, most of this association is due to underlying differences between children who move often and those who do not. [101] Residential mobility is also associated with poor neighbourhood quality. [102]

Childhood Maltreatment. Victims of child maltreatment engage in less outdoor active play. [103, 104] Childhood maltreatment (physical, sexual, and neglect) is linked to risky sexual behavior in female adolescents, an effect mediated by substance use and psychological distress. [105] Abuse is also linked to behaviors that indicate poor risk assessment such as increased gambling behavior. [106] Some evidence suggests that the rate of child abuse varies based on neighbourhood environment. [107]

	Outdoor Active Play	Health Risk Behaviour	Built Environment
Male Sex	+	++	N/A
Increased Age	– –	++	N/A
SES	–	++	++
Mental Health	– –	–	– –
Physical Health	–	– –	–
Residential Mobility	U	++	–
Ethnicity	Mix	Mix	Mix
Sleep	+	– –	–
Abuse	–	++	+
Rurality	+	++	+
Family Structure	–	Mix	U
Family Attachment	U	– –	U
Family Environment	+	–	U

U = unknown, – – = strong evidence for negative association, – = weak evidence for negative association, + weak evidence for positive association, ++ = strong evidence for positive association, Mix = mixed evidence.

Table 1: Known associations between outdoor active play, health risk behaviour, built environment, and covariates.

6 Additional Considerations

6.1 Overall Strengths and Limitations

This proposed thesis project proposes to examine the causal links from environment through outdoor active play to health risk behaviors. By accessing three different Canadian data sources, we hope to provide a balanced picture of this causal pathway. However, there are some limitations to what can be accomplished. The most recent longitudinal data source is more than 10 years old, and children may have different experiences now than they did in the past due in part to the increase in time spent online over this time period. Evaluation of the causal effects of neighbourhood-level exposures from observational data is difficult due to potential for bias due to self-selection into neighbourhoods. Lack of data on the amount of time spent in each activity necessitates assumptions in order to compare injury risks. The cross sectional nature of the data for manuscripts one and two also force certain assumptions. For manuscript one, we need to assume that the current state of the neighbourhood as observed represents the neighbourhood in the recent past when it was influencing outdoor active play. For manuscript two, we need to assume that previous injuries do not change the probability of injury for individuals.

6.2 Ethical considerations:

This proposed project will undergo ethics review by the Queen's University HSREB.

Informed Consent: *HBSC.* Informed consent is sought from both pupils and their parents or guardians (explicit or implicit, based on local school board custom), as participants are under the age of consent. [108] *NLSCY.* Explicit informed consent was obtained from the parent or legal guardian of all children enrolled in the study. *CHIRPP.* Consent is obtained from patients prior to data collection.

Harm Reduction: Care will be taken to prevent any individually identifiable data from being released. Data safety for the HBSC will be ensured by holding data on a password protected computer on site at Carruthers Hall, Queen's

University. All work with data from the NLSCY will be completed at the Queen's Research Data Centre. CHIRPP data will be held onsite at KGH on a hospital network computer.

Outputs: In order to ensure benefits are realized from this project, appropriate knowledge translation projects, including open access publications, infographics, interactive visualizations,[109] and/or guidelines for neighbourhood planning will be developed. If possible, a tool for calculating the playability index will be developed that can be used by any interested party. The specific nature of these projects and their target audience will depend on the findings.

6.3 Appropriateness for a PhD in Epidemiology.

Project Suitability. This project involves three different epidemiological study designs (one descriptive, one cross-sectional, one longitudinal), each with their own set of challenges. With respect to other methodological requirements, development of the classification algorithm for manuscript two, collection of neighbourhood features using GIS in manuscript one, and linking and managing the cycles of the NLSCY all represent significant time commitments. Data collection in the form of environmental data to inform the playability measure used in manuscript one and chart reviews for manuscript two will be completed by the student. Manuscripts one and three involve complex modelling challenges. In all manuscripts, care will be taken to use current epidemiological principles to maximize internal validity. The project scope has been approved at the outline stage of review.

Contributions of the Student. The project as proposed is to be completed primarily by the student with support from the supervisory team: Drs. William Pickett and Ian Janssen. Ideas for the manuscripts were developed by the State of Play team at UBC, by Drs. Pickett and Janssen, and by the student.

Feasibility: Access to the HBSC data will be facilitated by Drs. Pickett and Janssen, two of the Canadian HBSC investigators. Access to the NLSCY data will require a project proposal to be submitted to and approved by SSHRC; other Ph.D. students of Drs. Pickett and Janssen have previously gone through this process. Access to Kingston CHIRPP data will be obtained with help from Dr. Rob Brison and Dr. Chris Evans, colleagues of Dr. Pickett's. I have applied for funding to attend a workshop in Causal Inference from Observational data.

Item	Timeframe
Proposal Presentation	Fall 2018
Manuscript Preparation	Fall 2018 - Winter 2019
Writing Thesis	January 2020 - Summer 2020
Thesis Defence	August 2020

Table 2: Timeline of project completion.

References

[1] Amy D Curtis, Erica A Hinckson, and Tineke C A Water. Physical activity is not play: perceptions of children and parents from deprived areas. *N. Z. Med. J.*, 125(1365):38–47, November 2012.

[2] Patricia F Pearce, Joanne S Harrell, and Robert G McMurray. Middle-school children's understanding of physical activity: "if you're moving, you're doing physical activity". *J. Pediatr. Nurs.*, 23(3):169–182, 2008.

[3] J Pernsteiner and Johannes Pernsteiner. [eliminating free play hurts children. psychologists warn about full schedules and excessive caution]. *Kinderkrankenschwester*, 31(3):121, 2012.

[4] Arya Ansari, Kierra Pettit, and Elizabeth Gershoff. Combating obesity in head start: Outdoor play and change in children's body mass index. *J. Dev. Behav. Pediatr.*, 36(8):605–612, October 2015.

[5] Libby A Watson, Martyn C Baker, and Paul M Chadwick. Kids just wanna have fun: Children's experiences of a weight management programme. *Br. J. Health Psychol.*, 21(2):407–420, May 2016.

[6] Helena Elisabeth Elsje Caro, Teatske Maria Altenburg, Christine Dedding, and Mai Jeanette Maidy Chinapaw. Dutch primary schoolchildren's perspectives of Activity-Friendly school playgrounds: A participatory study. *Int. J. Environ. Res. Public Health*, 13(6), May 2016.

[7] Susan Herrington and Mariana Brussoni. Beyond physical activity: The importance of play and Nature-Based play spaces for children's health and development. *Curr. Obes. Rep.*, 4(4):477–483, 2015.

[8] Raktim Mitra, Guy E J Faulkner, Ron N Buliung, and Michelle R Stone. Do parental perceptions of the neighbourhood environment influence children's independent mobility? evidence from toronto, canada. *Urban Stud.*, 51(16):3401–3419, 2014.

[9] Jane E Barker, Andrei D Semenov, Laura Michaelson, Lindsay S Provan, Hannah R Snyder, and Yuko Munakata. Less-structured time in children's daily lives predicts self-directed executive functioning. *Frontiers in psychology*, 5:593, 2014.

[10] Marie-Jeanne Aarts, Wanda Wendel-Vos, Hans A M van Oers, Ien A M van de Goor, and Albertine J Schuit. Environmental determinants of outdoor play in children: a large-scale cross-sectional study. *Am. J. Prev. Med.*, 39(3):212–219, September 2010.

[11] Scyatta A Wallace, Torsten B Neilands, and Kathy Sanders Phillips. Neighborhood context, psychological outlook, and risk behaviors among urban african american youth. *Cultur. Divers. Ethnic Minor. Psychol.*, 23(1):59–69, January 2017.

[12] Jana Spilkova, Dagmar Dzúrova, and Michal Pitonak. Perception of neighborhood environment and health risk behaviors in prague's teenagers: a pilot study in a post-communist city. *Int. J. Health Geogr.*, 13:41, October 2014.

[13] Karen Glanz. *Health Behavior: Theory, Research, and Practice*. John Wiley & Sons, July 2015.

[14] Roger G Barker and Others. *Ecological psychology*, volume 7. Stanford University Press Stanford, CA, 1968.

[15] P H Moos. Social-ecological perspectives on health. *Health psychology: A handbook*, 1980.

[16] Urie Bronfenbrenner. Toward an experimental ecology of human development. *Am. Psychol.*, 32(7):513, 1977.

[17] Thomas A Glass and Matthew J McAtee. Behavioral science at the crossroads in public health: extending horizons, envisioning the future. *Soc. Sci. Med.*, 62(7):1650–1671, April 2006.

[18] Alice Moore and Helen Lynch. Accessibility and usability of playground environments for children under 12: A scoping review. *Scand. J. Occup. Ther.*, 22(5):331–344, 2015.

[19] Teun Remmers, Suzanne M L Broeren, Carry M Renders, Remy A Hirasing, Amy van Grieken, and Hein Raat. A longitudinal study of children's outside play using family environment and perceived physical environment as predictors. *Int. J. Behav. Nutr. Phys. Act.*, 11:76, June 2014.

[20] Marie-Jeanne Aarts, Sanne I de Vries, Hans A M van Oers, and Albertine J Schuit. Outdoor play among children in relation to neighborhood characteristics: a cross-sectional neighborhood observation study. *Int. J. Behav. Nutr. Phys. Act.*, 9(1):98, August 2012.

[21] Anna Timperio, Jacqueline Reid, and Jenny Veitch. Playability: Built and social environment features that promote physical activity within children. *Curr. Obes. Rep.*, 4(4):460–476, 2015.

[22] hillier-penn-hanson-grajewski-x.PDF.

[23] Valerie Carson, Andrei Rosu, and Ian Janssen. A cross-sectional study of the environment, physical activity, and screen time among young children and their parents. *BMC Public Health*, 14:61, January 2014.

[24] I H Akinci, R E Tarter, and L Kirisci. Concordance between verbal report and urine screen of recent marijuana use in adolescents. *Addict. Behav.*, 26(4):613–619, July 2001.

[25] D A Murphy, S Durako, L R Muenz, and C M Wilson. Marijuana use among HIV-positive and high-risk adolescents: a comparison of self-report through audio computer-assisted self-administered interviewing and urinalysis. *Am. J. Epidemiol.*, 152(9):805–813, November 2000.

[26] Nancy D Brener, John O G Billy, and William R Grady. Assessment of factors affecting the validity of self-reported health-risk behavior among adolescents: evidence from the scientific literature. *J. Adolesc. Health*, 33(6):436–457, December 2003.

[27] A Post, H Gilljam, I Rosendahl, L Meurling, S Bremberg, and M R Galanti. Validity of self reports in a cohort of swedish adolescent smokers and smokeless tobacco (snus) users. *Tob. Control*, 14(2):114–117, April 2005.

[28] Jonathan L Kwong, Don A Klinger, Ian Janssen, and William Pickett. Derivation of some contemporary scales to measure adolescent risk-taking in canada. *Int. J. Public Health*, October 2017.

[29] J Michael Oakes. The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Soc. Sci. Med.*, 58(10):1929–1952, May 2004.

[30] Randy Boyes, Dylan E. O'Sullivan, Brooke Linden, Michael McIsaac, and William Pickett. Gender-specific associations between involvement in team sport culture and canadian adolescents substance-use behavior. *SSM - Population Health*, 3:663 – 673, 2017.

[31] A T Galecki, T Burzykowski, S Chen, J A Faulkner, and J Ashton-Miller. Statistical power calculations for clustered continuous data. *Int. J. Knowl. Eng. Soft Data Paradig.*, 1(1):40–48, January 2009.

[32] Tyler J VanderWeele and Ilya Shpitser. A new criterion for confounder selection. *Biometrics*, 67(4):1406–1413, December 2011.

[33] Government of Canada, Statistics Canada. Archived – from urban areas to population centres. `https://www.statcan.gc.ca/eng/subjects/standard/sgc/notice/sgc-06`, February 2011. Accessed: 2018-8-8.

[34] Homan Lee, Katherine A Tamminen, Alexander M Clark, Linda Slater, John C Spence, and Nicholas L Holt. A meta-study of qualitative research examining determinants of children's independent active free play. *Int. J. Behav. Nutr. Phys. Act.*, 12:5, January 2015.

[35] Andrew N Hashikawa, Manya F Newton, Rebecca M Cunningham, and Martha W Stevens. Unintentional injuries in child care centers in the united states: a systematic review. *J. Child Health Care*, 19(1):93–105, 2015.

[36] Lina Gyllencreutz, Ewa Rolfsman, and Britt-Inger Saveman. Non-minor injuries among children sustained in an outdoor environment–a retrospective register study. *Int. J. Inj. Contr. Saf. Promot.*, 22(1):3–10, 2015.

[37] Peter Michael Klimek, David Juen, Enno Stranzinger, Rainer Wolf, and Theddy Slongo. Trampoline related injuries in children: risk factors and radiographic findings. *World J. Pediatr.*, 9(2):169–174, May 2013.

[38] David B Eager, Carl Scarrott, Jim Nixon, and Keith Alexander. Injury survey of a non-traditional 'soft-edged' trampoline designed to lower equipment hazards. *Int. J. Inj. Contr. Saf. Promot.*, 20(1):42–49, 2013.

[39] Stephanie Truelove, Leigh M Vanderloo, and Patricia Tucker. Defining and measuring active play among young children: a systematic review. *Journal of physical activity and health*, 14(2):155–166, 2017.

[40] Statistics Canada Government of Canada. 2016 census data tables – age (in single years) and average age (127) and sex (3) for the population of canada, census metropolitan areas, census agglomerations and census subdivisions, 2016 and 2011 censuses - 100% data. `http://www12.statcan.gc.ca/census-recensement/2016/dp-pd/dt-td/Rp-eng.cfm?TABID=6&LANG=E&APATH=3&DETAIL=0&DIM=0&FL=A&FREE=0&GC=0&GID=0&GK=0&GRP=1&PID=109524&PRID=10&PTYPE=109445&S=0&SHOWALL=0&SUB=0&Temporal=2016&THEME=115&VID=0&VNAMEE=&VNAMEF=&D1=0&D2=0&D3=0&D4=0&D5=0&D6=0`, May 2017. Accessed: 2018-4-4.

[41] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3(Jan):993–1022, 2003.

[42] Khairun-Nisa Hassanali, Yang Liu, and Thamar Solorio. Using latent dirichlet allocation for child narrative analysis. In *Proceedings of the 2013 Workshop on Biomedical Natural Language Processing*, pages 111–115, 2013.

[43] Raphael Cohen, Iddo Aviram, Michael Elhadad, and Noémie Elhadad. Redundancy-aware topic modeling for patient record notes. *PLoS One*, 9(2):e87555, February 2014.

[44] Liraz Fridman, Jessica L Fraser-Thomas, Steven R McFaull, and Alison K Macpherson. Epidemiology of sports-related injuries in children and youth presenting to canadian emergency departments from 2007-2010. *BMC Sports Sci. Med. Rehabil.*, 5(1):30, December 2013.

[45] William Pickett, Susan Streight, Kelly Simpson, and Robert J Brison. Injuries experienced by infant children: a population-based epidemiological analysis. *Pediatrics*, 111(4 Pt 1):e365–70, April 2003.

[46] Statistics Canada Government of Canada. Statistics canada: 2006 community profiles. March 2007.

[47] Nandini Mundkur. Neuroplasticity in children. *Indian J. Pediatr.*, 72(10):855–857, October 2005.

[48] Natasha Noble, Christine Paul, Heidi Turon, and Christopher Oldmeadow. Which modifiable health risk behaviours are related? a systematic review of the clustering of smoking, nutrition, alcohol and physical activity ('SNAP') health risk factors. *Prev. Med.*, 81:16–41, December 2015.

[49] R Jessor. Problem-behavior theory, psychosocial development, and adolescent problem drinking. *Br. J. Addict.*, 82(4):331–342, April 1987.

[50] Michael Mobley and Heejung Chun. Testing jessor's problem behavior theory and syndrome: a nationally representative comparative sample of latino and african american adolescents. *Cultur. Divers. Ethnic Minor. Psychol.*, 19(2):190–199, April 2013.

[51] Elizabeth T Couch, Ellen Darius, Margaret M Walsh, and Benjamin W Chaffee. Smokeless tobacco Decision-Making among rural adolescent males in california. *J. Community Health*, 42(3):544–550, June 2017.

[52] Jennifer B Unger, Lourdes Baezconde-Garbanati, Anamara Ritt-Olson, Daniel W Soto, and Chih-Ping Chou. Predictors of growth trajectories of substance use from 9th to 11th grade among hispanic adolescents in southern california. In *Drug Use Trajectories Among Minority Youth*, pages 151–169. Springer, Dordrecht, 2016.

[53] Rebecca J Johnson, Kevin D McCaul, and William M P Klein. Risk involvement and risk perception among adolescents and young adults. *J. Behav. Med.*, 25(1):67–82, February 2002.

[54] Statistics Canada Government of Canada. Surveys and statistical programs - national longitudinal survey of children and youth (NLSCY). `http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=4450#a4`, May 2009. Accessed: 2018-6-21.

[55] W Pickett, S Dostaler, W Craig, I Janssen, K Simpson, S Danielle Shelley, and W F Boyce. Associations between risk behavior and injury and the protective roles of social environments: an analysis of 7235 canadian school children. *Inj. Prev.*, 12(2):87–92, April 2006.

[56] R Koven, M A McColl, P Ellis, and W Pickett. Multiple risk behaviour and its association with head and neck injuries: a national analysis of young canadians. *Prev. Med.*, 41(1):240–246, July 2005.

[57] Ruth H Keogh, Rhian M Daniel, Tyler J VanderWeele, and Stijn Vansteelandt. Analysis of longitudinal studies with repeated outcome measures: Adjusting for Time-Dependent confounding using conventional methods. *Am. J. Epidemiol.*, 187(5):1085–1092, May 2018.

[58] Andrew Bell and Kelvyn Jones. Age, period and cohort processes in longitudinal and life course analysis: A multilevel perspective. In Claudine Burton-Jeangros, Stéphane Cullati, Amanda Sacker, and David Blane, editors, *A Life Course Perspective on Health Trajectories and Transitions*. Springer, Cham (CH), September 2016.

[59] Miguel Hernan and James Robins. *Causal Inference*.

[60] Maya L Petersen, Kristin E Porter, Susan Gruber, Yue Wang, and Mark J van der Laan. Diagnosing and responding to violations in the positivity assumption. *Stat. Methods Med. Res.*, 21(1):31–54, October 2010.

[61] Microdata user's guide.

[62] Randi Melissa Schuster, Robin Mermelstein, and Laurie Wakschlag. Gender-specific relationships between depressive symptoms, marijuana use, parental communication and risky sexual behavior in adolescence. *J. Youth Adolesc.*, 42(8):1194–1209, August 2013.

[63] Jason N Bocarro, Myron F Floyd, William R Smith, Michael B Edwards, Courtney L Schultz, Perver Baran, Robin A Moore, Nilda Cosco, and Luis J Suau. Social and environmental factors related to boys' and girls' Park-Based physical activity. *Prev. Chronic Dis.*, 12:E97, June 2015.

[64] Andrew T Kaczynski, Sonja A Wilhelm Stanis, Gina M Besenyi, and Stephanie Child. Differences in youth and adult physical activity in park settings by sex and race/ethnicity. *Prev. Chronic Dis.*, 10:E42, 2013.

[65] Ellen Haug, Torbjørn Torsheim, James F Sallis, and Oddrun Samdal. The characteristics of the outdoor school environment associated with physical activity. *Health Educ. Res.*, 25(2):248–256, April 2010.

[66] Charlotte Skau Pawlowski, Henriette Bondo Andersen, Jens Troelsen, and Jasper Schipperijn. Children's physical activity behavior during school recess: A pilot study using GPS, accelerometer, participant observation, and Go-Along interview. *PLoS One*, 11(2):e0148786, February 2016.

[67] Bettina Bringolf-Isler, Leticia Grize, Urs Mäder, Nicole Ruch, Felix H Sennhauser, Charlotte Braun-Fahrländer, and SCARPOL team. Built environment, parents' perception, and children's vigorous outdoor play. *Prev. Med.*, 50(5-6):251–256, May 2010.

[68] Shoo Thien Lee, Jyh Eiin Wong, Wei Wen Ong, Mohd Noor Ismail, Paul Deurenberg, and Bee Koon Poh. Physical activity pattern of malaysian preschoolers: Environment, barriers, and motivators for active play. *Asia. Pac. J. Public Health*, 28(5 Suppl):21S–34S, July 2016.

[69] Jessica L Simmons, James P Whelan, Andrew W Meyers, and Emerson M Wickwire. Gambling outcome expectancies and gambling behavior among African-American adolescents: Gender as a moderating variable. *J. Gambl. Stud.*, 32(1):205–215, March 2016.

[70] Robert J Wellman, Erika N Dugas, Hartley Dutczak, Erin K O'Loughlin, Geetanjali D Datta, Beatrice Lauzon, and Jennifer O'Loughlin. Predictors of the onset of cigarette smoking: A systematic review of longitudinal Population-Based studies in youth. *Am. J. Prev. Med.*, 51(5):767–778, 2016.

[71] F T Filippidis, I T Agaku, and C I Vardavas. Geographic variation and socio-demographic determinants of the co-occurrence of risky health behaviours in 27 european union member states. *J. Public Health*, 38(2):e13–20, June 2016.

[72] Leonie Uijtdewilligen, Joske Nauta, Amika S Singh, Willem van Mechelen, Jos W R Twisk, Klazine van der Horst, and Mai J M Chinapaw. Determinants of physical activity and sedentary behaviour in young people: a review and quality synthesis of prospective studies. *Br. J. Sports Med.*, 45(11):896–905, 2011.

[73] R Andrew Chambers, Jane R Taylor, and Marc N Potenza. Developmental neurocircuitry of motivation in adolescence: a critical period of addiction vulnerability. *Am. J. Psychiatry*, 160(6):1041–1052, June 2003.

[74] Niina Kolehmainen, Craig Ramsay, Lorna McKee, Cheryl Missiuna, Christine Owen, and Jill Francis. Participation in physical play and leisure in children with motor impairments: Mixed-Methods study to generate evidence for developing an intervention. *Phys. Ther.*, 95(10):1374–1386, October 2015.

[75] Nicola McLaren, Steven J Kamper, Rebecca Hodder, John Wiggers, Luke Wolfenden, Jennifer Bowman, Elizabeth Campbell, Julia Dray, and Christopher M Williams. Increased substance use and poorer mental health in adolescents with problematic musculoskeletal pain. *J. Orthop. Sports Phys. Ther.*, 47(10):705–711, October 2017.

[76] Ruopeng An, Yan Yang, and Kaigang Li. Residential neighborhood amenities and physical activity among U.S. children with special health care needs. *Matern. Child Health J.*, 21(5):1026–1036, May 2017.

[77] Sara K Crawford, Karen N Stafford, Sarah M Phillips, Kathleen J Scott, and Patricia Tucker. Strategies for inclusion in play among children with physical disabilities in childcare centers: an integrative review. *Phys. Occup. Ther. Pediatr.*, 34(4):404–423, 2014.

[78] Elizabeth A Courtney-Long, Sebastian D Romano, Dianna D Carroll, and Michael H Fox. Socioeconomic factors at the intersection of race and ethnicity influencing health risks for people with disabilities. *J Racial Ethn Health Disparities*, 4(2):213–222, April 2017.

[79] [PDF]Urban and rural student substance use - canadian centre on ..

[80] Carolyn McAndrews, Kirsten Beyer, Clare E Guse, and Peter Layde. How do the definitions of urban and rural matter for transportation safety? re-interpreting transportation fatalities as an outcome of regional development processes. *Accid. Anal. Prev.*, 97:231–241, December 2016.

[81] Maria M Galano, Margaret D McGuire, Andrew Grogan-Kaylor, Nora Montalvo-Liendo, and Sandra A Graham-Bermann. Ethno-Racial differences in children's attitudes and beliefs about violence. *Violence Vict.*, 32(3):452–465, June 2017.

[82] J Salmon, Tremblay, S J Marshall, C Hume, Jo Salmon, Mark S Tremblay, Simon J Marshall, and Clare Hume. Health risks, correlates, and interventions to reduce sedentary behavior in young people. *Am. J. Prev. Med.*, 41(2):197–206, 2011.

[83] Regina A Shih, Jeremy N V Miles, Joan S Tucker, Annie J Zhou, and Elizabeth J D'Amico. Racial/ethnic differences in adolescent substance use: mediation by individual, family, and school factors. *J. Stud. Alcohol Drugs*, 71(5):640–651, September 2010.

[84] Clement J Bottino, Sheryl L Rifas-Shiman, Ken P Kleinman, Emily Oken, Susan Redline, Diane Gold, Joel Schwartz, Steven J Melly, Petros Koutrakis, Matthew W Gillman, and Elsie M Taveras. The association of urbanicity with infant sleep duration. *Health Place*, 18(5):1000–1005, September 2012.

[85] Megan E Patrick, Patrick Wightman, Robert F Schoeni, and John E Schulenberg. Socioeconomic status and substance use among young adults: a comparison across constructs and drugs. *J. Stud. Alcohol Drugs*, 73(5):772–782, September 2012.

[86] Gavin R Jenkins, Hon K Yuen, Emily J Rose, Amy I Maher, Kristina C Gregory, and Megan E Cotton. Disparities in quality of park play spaces between two cities with diverse income and Race/Ethnicity composition: A pilot study. *Int. J. Environ. Res. Public Health*, 12(7):8009–8022, July 2015.

[87] Amiri Moussa, Soori Hamid, Ainy Elaheh, and Mehmandar Mohammad Reza. Parent's and children's judgements about their outdoor environment in relation to children's injuries. *J. Pak. Med. Assoc.*, 63(12):1504–1508, December 2013.

[88] J S Kjeldsen, M F Hjorth, R Andersen, K F Michaelsen, I Tetens, A Astrup, J-P Chaput, and A Sjödin. Short sleep duration and large variability in sleep duration are independently associated with dietary risk factors for obesity in danish school children. *Int. J. Obes.*, 38(1):32–39, January 2014.

[89] David A Reichenberger, Clayton J Hilmert, Leah A Irish, Molly Secor-Turner, and Brandy A Randall. Associations between sleep and Health-Risk behaviors in a rural adolescent population. *J. Pediatr. Health Care*, 30(4):317–322, July 2016.

[90] Thomas A Arcury, Cynthia K Suerken, Edward H Ip, Justin B Moore, and Sara A Quandt. Residential environment for outdoor play among children in latino farmworker families. *J. Immigr. Minor. Health*, 19(2):267–274, April 2017.

[91] S G O'Connor, J P Maher, B R Belcher, A M Leventhal, G Margolin, E T Shonkoff, and G F Dunton. Associations of maternal stress with children's weight-related behaviours: a systematic literature review. *Obes. Rev.*, 18(5):514–525, 2017.

[92] Roger Kobak, Kristyn Zajac, Caroline Abbott, Abby Zisk, and Nadia Bounoua. Atypical dimensions of caregiver-adolescent interaction in an economically disadvantaged sample. *Dev. Psychopathol.*, 29(2):405–416, May 2017.

[93] Jillian Lee Wiggins, Colter Mitchell, Luke W Hyde, and Christopher S Monk. Identifying early pathways of risk and resilience: The codevelopment of internalizing and externalizing symptoms and the role of harsh parenting. *Dev. Psychopathol.*, 27(4 Pt 1):1295–1312, November 2015.

[94] Diana R Samek, Martha A Rueter, Margaret A Keyes, Matt McGue, and William G Iacono. Parent involvement, sibling companionship, and adolescent substance use: A longitudinal, genetically informed design. *J. Fam. Psychol.*, 29(4):614–623, August 2015.

[95] David Córdova, Seth J Schwartz, Jennifer B Unger, Lourdes Baezconde-Garbanati, Juan A Villamar, Daniel W Soto, Sabrina E Des Rosiers, Tae Kyoung Lee, Alan Meca, Miguel Ángel Cano, Elma I Lorenzo-Blanco, Assaf Oshri, Christopher P Salas-Wright, Brandy Piña-Watson, and Andrea J Romero. A longitudinal test of the Parent-Adolescent family functioning discrepancy hypothesis: A trend toward increased HIV risk behaviors among immigrant hispanic adolescents. *J. Youth Adolesc.*, 45(10):2164–2177, October 2016.

[96] Huilan Xu, Li Ming Wen, Louise L Hardy, and Chris Rissel. Mothers' perceived neighbourhood environment and outdoor play of 2- to 3.5-Year-Old children: Findings from the healthy beginnings trial. *Int. J. Environ. Res. Public Health*, 14(9), September 2017.

[97] Shannon D Boone and Amy M Brausch. Physical activity, exercise motivations, depression, and nonsuicidal Self-Injury in youth. *Suicide Life Threat. Behav.*, 46(5):625–633, October 2016.

[98] Kyoung-Bok Min, Hyun-Jin Kim, Hye-Jin Kim, and Jin-Young Min. Parks and green areas and the risk for depression and suicidal indicators. *Int. J. Public Health*, 62(6):647–656, July 2017.

[99] Regina Reklaitiene, Regina Grazuleviciene, Audrius Dedele, Dalia Virviciute, Jone Vensloviene, Abdonas Tamosiunas, Migle Baceviciene, Dalia Luksiene, Laura Sapranaviciute-Zabazlajeva, Ricardas Radisauskas, Gailute Bernotiene, Martin Bobak, and Mark J Nieuwenhuijsen. The relationship of green space, depressive symptoms and perceived general health in urban population. *Scand. J. Public Health*, 42(7):669–676, November 2014.

[100] D Brown, M Benzeval, V Gayle, S Macintyre, D O'Reilly, and A H Leyland. Childhood residential mobility and health in late adolescence and adulthood: findings from the west of scotland twenty-07 study. *J. Epidemiol. Community Health*, 66(10):942–950, October 2012.

[101] Tim Morris, David Manley, Kate Northstone, and Clive E Sabel. On the move: Exploring the impact of residential mobility on cannabis use. *Soc. Sci. Med.*, 168:239–248, November 2016.

[102] Ade Kearns and Phil Mason. Regeneration, relocation and health behaviours in deprived communities. *Health Place*, 32:43–58, March 2015.

[103] D Darwish, G B Esquivel, J C Houtz, and V C Alfonso. Play and social skills in maltreated and non-maltreated preschoolers during peer interactions. *Child Abuse Negl.*, 25(1):13–31, January 2001.

[104] A C Howard and A C Howard. Developmental play ages of physically abused and nonabused children.[erratum appears in am J occup ther 1987 jan;41(1):27]. *Am. J. Occup. Ther.*, 40(10):691–695, 1986.

[105] Kristen Clements-Nolle, Sandra Larson, Aliya Buttar, and Lindsey Dermid-Gray. Childhood maltreatment and unprotected sex among female juvenile offenders: Evidence of mediation by substance abuse and psychological distress. *Womens. Health Issues*, 27(2):188–195, March 2017.

[106] Wendy Lane, Paul Sacco, Katherine Downton, Emilie Ludeman, Lauren Levy, and J Kathleen Tracy. Child maltreatment and problem gambling: A systematic review. *Child Abuse Negl.*, 58:24–38, 2016.

[107] Bridget Freisthler, Paul J Gruenewald, Lori Ring, and Elizabeth A LaScala. An ecological assessment of the population and environmental correlates of childhood accident, assault, and child abuse injuries. *Alcohol. Clin. Exp. Res.*, 32(11):1969–1975, November 2008.

[108] Health behaviour in school-aged children (HBSC) study protocol 2009/10.

[109] Explorable explanations. `http://explorabl.es/`. Accessed: 2017-11-27.

7 Appendix A: Comments on Suggested Improvements

Outline Feedback

Comment 1. Suitability as a PhD in Epidemiology. The committee agreed that the project as proposed: (1) would be suitable in terms of the amount of work (if anything, too ambitious; (2) was cohesive and fit together thematically; (3) involved multiple datasets with different challenges and opportunities; (4) had depth of statistical and methodological complexity consistent with what was required in our department.

Response 1. Changes have been made to the scope - reduced number of years in the CHIRPP study and focused the definition of risk behavior.

Comment 2. Conceptual Framework. There is a need for a more comprehensive conceptual framework that builds on the existing figure, and considers risk behaviours, healthy behaviours, and other covariates, and directs the reader to which of these will be a focus.

Response 2. Figures have been added to this effect (Figures 1 and 3)

Comment 3. Rates (manuscript 1 [now 2]). There needs to be a clear and informed section of the proposal that addresses how the denominators will be created for the Kingston area hospitals, so that rates can be calculated.

Response 3. This section has been added (Section REF)

Comment 4. Scope (manuscripts 1,3). There were concerns expressed around the scope of the manuscripts, and specifically: (a) Manuscript 1; 20 years of data; 4 types of activity leading to injury; many covariates; and the recommendation was to make some practical decisions so that the scope for the data collection, analysis and its interpretation could be done in one paper; (b) Manuscript 3; the scope and practicality of measuring the playability index for all participants was questioned. It was recommended that we need to consider the feasibility of having the measured required for the playability index via an automated process, and have multiple options considered in the proposal should the preferred option not be feasible. (c) there was concern that Randy examines the scope of the entire project, and be prepared to make some decisions to cut if it is not feasible to complete within the remaining PHD time (2 and a bit years).

Response 4. See point 1.

Comment 5. Functional Logistic Regression. Recommendation is that the statistical approach used for the longitudinal analysis may not be optimal, and Randy should examine other options (e.g., logistic or another regression method, with conventional approaches to time-dependent covariate data). Need to also consider the timing of the measures and developmental stages of children in making these decisions.

Response 5. The suggestion of functional logistic regression has been removed. A discussion of the particular demands of this project and expected obstacles has been added.

Comment 6. Ensure that writing is polished, and tells a cohesive and progressive story, and that the objectives and hypotheses are explicit.