Title

The Built Environment and Physical Activity: Empirical Methods and Data Resource

Author

Boarnet, Marlon G.

Publication Date

2003-11-01

The Built Environment and Physical Activity:
Empirical Methods and Data Resources

Marlon G. Boarnet
Department of Planning, Policy, and Design
Department of Economics
and Institute of Transportation Studies
University of California, Irvine

Paper prepared for the Transportation Research Board and the Institute of Medicine
Committee on Physical Activity, Health, Transportation, and Land Use

November 17, 2003

DRAFT
DO NOT QUOTE OR CITE WITHOUT PERMISSION OF THE AUTHOR

The author thanks Raquel Girvin for research assistance. The author bears full responsibility for any errors or omissions.

I. Introduction

Does a person's environment influence their physical activity? Intuition, theory, and preliminary evidence all suggest that there is an association between environment and physical activity, but questions of causality and magnitude remain poorly answered, in large part due to data challenges. If public health policy is to make meaningful links to the built environment, the literature will require careful tests of causal links and an understanding of the magnitude of those links. This paper reviews the data that are available for testing hypotheses about the built environment, physical activity, and health outcomes, both to educate the research community about existing data and current challenges and to illuminate data gaps that should be addressed as this research agenda moves forward.[1]

The hypothesis that much of the emerging public health and health promotion literature seeks to test is that the built environment influences physical activity in ways that improve individual health. The latter causal chain, from physical activity to health, is well understood, and the epidemiological literature gives strong evidence of links between physical activity, mortality, and specific diseases (e.g., Pate et al., 1995; Paffenbarger et al., 1986; Leon et al, 1987; Ekelund et al., 1988; Blair et al, 1989; Morris et al., 1990; Sandvik et al., 1993; U.S. Department of Health and Human Services, 1996). For that reason, we focus here on data that can support tests of the influence of the built environment on physical activity. An ideal data set and empirical test would examine the causal path from the built environment to physical activity to health outcomes, so data that can support such a test will receive attention, but the first of those two causal chains, from the built environment to physical activity, poses special data problems and, largely for that reason, is the less well understood of the two causal links. Hence data that can test hypotheses about the built environment and physical activity are the primary focus of this paper.

The data needed to understand how the built environment influences physical activity are currently spread across different data sources, in different fields that have asked different research questions. Physical activity data are most well developed in the fields of health promotion and public health. The research on physical activity in those fields has traditionally focused more on individual psychological or motivational factors, family or peer influences, or programmatic interventions. The questions asked were often not geographic in nature, and for that reason information that allows a link to physical geography, and through that to measures of the built environment, are not prominent in much of the health-based literature. Conversely, literatures that have made the most progress in measuring the built environment – elements of planning, design-behavior, or

[1] The focus of this paper is on the built environment. The natural environment likely influences physical activity, and research has examined the links between trails, nature, and physical activity (cites). Yet we note that, in 1990, 75% of Americans lived in census defined urbanized areas, and even in rural locales a person's experience with the environment on a daily basis is largely with the built environment. Our focus here is largely on built environments that exist in small towns, communities, and urban areas. This does not preclude the natural environment as a determinant of physical activity, but for purposes of the data surveyed here the emphasis is on links to the built environment.

applied geography research – have often not addressed physical activity. Hence the measures of the built environment from those literatures are typically not tailored to illuminate physical activity patterns. One particular challenge is that the standard of geographic precision in measuring the built environment – a street address or a similarly precise location defined by map coordinates – clashes with concerns about subject anonymity in health-based physical activity surveys. Some health surveys gather sensitive data, and so provide survey subjects strong assurances of anonymity. While address-based data need not compromise that anonymity, there are differences in research culture and issues in research management that need to be addressed to allow both the fine spatial detail required to measure the built environment while maintaining guarantees of subject anonymity. All of these challenges can be overcome. The data tools available allow a rich set of information on health outcomes, physical activity, and the built environment. The task now is to put these different pieces together in ways that illuminate whether and how the built environment influences physical activity and health.

This paper groups the existing data sources into three areas – (1) data on physical activity or health outcomes, (2) data on the built environment, and (3) data from studies of travel behavior and the built environment. The first two types of data are fundamental to the study of physical activity and the built environment. The last group of data, drawn from studies of travel behavior and the built environment, illuminate many of the issues inherent in linking an outcome variable to the environment. In that last group, we will focus on studies of non-motorized travel behavior and the built environment, since walking and bicycling are common forms of physical activity.

II. Research Design

Simply canvassing data sources is not sufficient to advance research in this area. Examining data in an atheoretical or empiricist manner can establish associations, but not causal links. Studies of aggregate data have established intriguing evidence that the built environment is associated with patterns of physical activity (cites), but evidence of causality is needed for policy intervention.

Establishing causal links requires two touchstones – individual data and theoretically informed empirical tests. Aggregate data will rarely illuminate behavioral links as well as individual data can, and that general point holds when studying the built environment and physical activity. Several fields offer insights into studies of the built environment and physical activity. Those fields include health promotion, motivational determinants of physical activity, and, for non-motorized travel, theories of travel behavior. The insights from these theories focus either on individual choice behavior, often assuming that individuals are rational actors who either optimize or satisfice, or on the cognitive processes underlying choice decisions. Some of the available theories have gaps when applied to the full range of physical activity. As an example, theories of travel behavior typically assume that travel is a derived demand – that persons travel not for the pleasure of travel itself, but to consume goods or activities required by travel.[2] While that might

[2] For a discussion, see, e.g., Handy, Boarnet, Ewing, and Killingsworth (2002) or Small (1992), and for a critique of this approach, see Mohktarian (year).

be approximately correct as a theory of travel behavior, it has obvious shortcomings explaining leisure time physical activity that is often pursued as an end in itself. Alternatively, theories of physical activity often focus on cognitive processes and motivation, and give comparatively little attention to the built environment (see, e.g., the discussion in King, et al., 2002).

Other authors have reviewed various theories about the link between the built environment and physical activity (e.g. King, Stokols, Talen, Brassington, and Killingsworth, 2002), so our focus here is on elements of research design that inform the selection of data sources. There are three research design considerations that are fundamental in studying the built environment and physical activity, and that cannot be separated from a discussion of data.

1. Data should allow researchers to separate the causal influence of the built environment on physical activity from the reverse causal influence that persons might choose locations (especially residential locations) based on their desired level of physical activity.

2. Data should be longitudinal, when possible, to allow both long-term tracking of links between the built environment, physical activity, and health outcomes and to help isolate the independent effect of the built environment by allowing researchers to assess the simultaneous relationship between physical activity and an individual's choice of neighborhood environment. Note that cross-sectional data will also be useful, and we do not call for an exclusive focus on longitudinal data.

3. Data should allow multiple levels of geographic detail, with special focus on fine-grained, neighborhood level built environment characteristics likely to influence physical activity.

Each of these issues is discussed below.

A. Residential Selection

Persons might choose their environment in part based on their desired level of physical activity. It does not take much imagination to believe that an avid surfer would choose to live near the beach, or that a ski enthusiastic would move near the mountains. Generalizing to other, more common forms of physical activity, do persons who wish to walk choose residences in pedestrian-oriented neighborhoods near parks? If so, the associations between physical activity and urban form might represent persons' residential location choice rather than an influence of the built environment on activity.[3]

---

[3] A recent article, by Ewing et al. (2003), illustrates some of the difficulties. Ewing et al, in a national study, find that individual physical activity and body-mass index are positively associated with a measure of urban sprawl in their county of residence. One critique would be that persons might choose their county of residence based in part on their desired activity pattern. Yet counties are somewhat large, and it is reasonable to suspect that the factors that influence a choice of county, as opposed to a choice of

Yet one might also reasonably suppose that the happenstance of living near a beach, mountain, park, or walking-friendly environment would influence a person's physical activity. The issue is disentangling what is almost certainly two-way causality, and attributing the independent and causal influence of the built environment on physical activity in a way that can inform planning and public health policy.

This problem has surfaced in other realms. For example, the literature on travel behavior and the built environment has noted that persons might choose residential locations in part based on how they wish to travel.[4] Boarnet and Sarmiento (1998) used an instrumental variables technique to control for residential location choice, and Krizek (2003) used data on persons who moved to associate the resulting changes in urban form with changes in travel behavior. Both are among the possible solutions to the similar problem in the context of physical activity research. Those and other techniques are discussed below.

With cross-sectional data on individuals, a researcher could model both an individual's amount of physical activity and their residential location choice decision. This could be either a fully specified joint (two-equation) model of residential location choice and physical activity, or what econometricians call a reduced form that examines only physical activity, while controlling for factors that influence residential location choice.[5] In either method, the data should include information that inform where a person might live, and physical activity data sets will often have a limited set of information that could inform individual location choice decisions. Still, the matter is not futile, since variables that influence location choice include the basic demographic variables that are tracked in many physical activity studies – age, occupation, race, income, marital status, family size, and the like. For an example drawn from the study of the built environment and travel behavior, see Boarnet and Sarmiento (1998). Implementing such a technique would involve regression techniques that can handle multiple endogenous variables (see, e.g., Johnston and Dinardo, 1995). In terms of data, these techniques require that the researcher has some variables that influence housing location choice that do not influence the choice of physical activity.[6] Because many of the sociodemographic characteristics

---

neighborhood, have more to do with job opportunities, family ties, and the like than with desired physical activity. If one believes that persons do not choose their county based on their desired activity patterns, this is evidence of a causal influence from the built environment to physical activity, and also to body mass index. Yet the coarse geographic scale of the research, while lessening questions of two-way causality, also limit the policy insight. Many policy suggestions relate to the design of neighborhoods as opposed to entire counties, which are often amalgams of a broad spectrum of types of built environments and neighborhoods. Hence the method that Ewing et al. (2003) used to lessen dual causality problems in their study also reduces the direct policy applicability of the results, especially as that would relate to theories of neighborhood design.

[4] The problem is more general, as persons might choose several locations – home, work, school, etc. – based in part on how they wish to travel. A similar generalization might apply to physical activity, and so the choice of residential, work, school, and other environments should be considered. Here we only focus on the choice of residential environments.

[5] For studies of individual location choice, see, e.g., Quigley (), Linneman (), Levine ().

[6] More technically, the condition that must be satisfied is that variables that are correlated with residential or neighborhood location choice but not correlated with the error term in a regression for physical activity are available.

that influence housing location or neighborhood choice likely also influence an individual's choice of physical activity, some care must be taken when specifying the regression models.

An alternative strategy is to observe when a person moves and to draw associations between changes in the built environment at the old residential location versus the new residential location and to also observe changes in physical activity. As mentioned above, Krizek (2002) implemented this strategy to examine the influence of the built environment on travel behavior. Yet moving is often associated with a large number of life changes; marital status, job, family size, employment status, and income often change coincident with a move. Studies that correlate residential moves with changes in the physical activity should take care to be certain that the built environment, as opposed to several other likely coincident life changes, is influencing any observed changes in physical activity.

A third strategy is to study children, on the assumption that children do not choose their residential location, and hence the built environment around their residence can be viewed as exogenous to their desired level of physical activity. This strategy has been used in research on access to employment and labor market success. The issue in that literature is analogous to what we are discussing here – persons who have a tenuous association with the labor market might choose residential locations with poor access to jobs, but the poor access (i.e. being distant from job centers) may not cause an individual's labor market outcomes. In this line of research – called the spatial mismatch literature among economists and geographers – children have been studied because their residential location is chosen by their parents, at least in most instances. An early study that examined children, specifically teens, in the spatial mismatch literature is Ellwood (1986), and O'Regan and Quigley (1998) summarize several other studies. One important difficulty of particular concern in the realm of physical activity research is that parents might impart attitudes about physical activity to their children, and if those parents choose residential locations based on their desired level of physical activity, correlations between the built environment and children's level of physical activity may not demonstrate a causal effect of the built environment.

A fourth method of overcoming complications from residential selection is to look for natural experiments that reproduce, at least to some extent, the random assignment of subjects to experimental and control groups that is typical of laboratory research. While such "natural experiments" or "quasi-experiments" have become increasingly popular research tools in the behavioral sciences during the past three decades, the prospects for using natural experiments in the context of the built environment are limited. Changes in built environments are rarely exogenous to residential location choices. The most significant built environment changes are typically developments or redevelopments that are large enough to likely induce some residential relocation, and hence reduce the value of the natural or quasi-experiment.

Having said that, some opportunities for natural experiments in the study of the built environment and physical activity do exist. For example, Boarnet, Day, Anderson, and

McMillan, with various other researchers, are studying construction projects designed to make walking and bicycling near California schools safer. That evaluation of the California Safe Routes to School construction program suggests that the intervention has had some impact on walking behavior near selected school sites (California Safe Routes to Schools, December, 2003). Other opportunities for quasi-experiments might exist. For example, differences in infrastructure, park, or school financing or site selection practices across states might create opportunities for natural experiments if such practices lead to systematic differences in built environments across metropolitan areas. This assumes that persons choose metropolitan areas based on factors other than their desired level of physical activity – an assumption that may be accurate. Yet overall, insights into the two-way causality between the built environment and physical activity are more likely to come from data that are complete enough to allow an examination of both residential location choices and physical activity, rather than from natural or quasi-experiments.

B. Longitudinal Data

Longitudinal data on physical activity and health that can be linked to the built environment can be useful for two reasons. First, such data allow a more complete examination of both residential location choices and physical activity. Second, longitudinal data can be used to track changes in physical activity over a life course.

Dealing with the first advantage, longitudinal data provide an opportunity to observe how an individual's physical activity changes when their residential location, and hence built environment, change. For reasons noted above, such associations will not determine causality, in large part because of the many other life changes that are coincident with changes in residential location. Still, longitudinal data give researchers more ability to examine the determinants of both residential location and physical activity, and so provide a better platform for examining these issues that similar data available only in a cross-section.

In terms of the second issue, longitudinal data allow tracking physical activity outcomes over a life course. Do physical activity habits acquired in youth persist into adulthood, implying that built environment interventions for youth might be more important? More generally, how does any influence of the built environment persist over both the life course and changes in residential location? Longitudinal data on both health outcomes and physical activity allow researchers to examine how physical activity patterns earlier in life correlate with health outcomes in later years. Some of these questions have already been studied in the epidemiological literature (check), but longitudinal data can allow a more comprehensive focus on the influence of the built environment over the life course.

C. Geographic Scale

Geography, and hence geographic scale, is fundamental to studies that attempt to link behavior to the built environment. Handy (1993) discussed this in relation so land use –

travel behavior interactions, noting that some forms of travel are influenced by the regional transportation network, while others are influenced by the more local, or neighborhood transportation network. When used in studies of land use – travel behavior links, regional typically is a scale of a few miles, corresponding to average commute distances or market areas for regional shopping malls, while local travel is typically within distances of less than a mile, and in some cases less than a quarter of a mile. Yet these definitions are loose, and measuring the built environment at different geographic scales has so far been limited by data availability and commonly accepted rules of thumb (see, e.g., Boarnet and Sarmiento, 1998 or Greenwald and Boarnet, 2001).

For physical activity, the geography in a local environment, or neighborhood, is likely the most important determinant, and much of the discussion of the built environment in this paper focuses on neighborhood-scale measures. Yet that still gives imprecise guidance on the appropriate geographic scale that corresponds to a neighborhood. Transportation studies have typically used a quarter-mile radius to demarcate walking-oriented neighborhoods, based on evidence that a quarter-mile is in many cases a reasonable limit to how far persons will walk in settings that provide alternatives (Untermann, year). Yet whether that scale is linked to all forms of physical activity is still unclear. In general, the issue of geographic scale is under-examined in recent literatures that relate to behavior to the built environment. For data sources, this implies that data that allow varying scales of geography are to be preferred.

[To insert later: Primer on census geography and geographic scale]

III. Physical Activity Data

Measuring physical activity can be complex. A common measure is energy expenditure, but other aspects of activity, including aerobic intensity and strength, also relate to health outcomes (Kriska and Casperson, 1997). Here we will focus on methods that measure energy expenditure.

Measuring energy expenditure can be accomplished either by methods that directly measure metabolic energy rate, or methods that measure the amount, duration, and intensity of activity, sometimes for purposes of converting that measure into a metabolic expenditure rate. Even here, matters can get more complex. Metabolic energy expenditure is typically measure in units of METs, where one MET is basal (resting) metabolic energy expenditure, set at 3.5 ml of oxygen consumed per kilogram of body mass per minute (Kriska and Casperson, 1997). Because energy expenditure differs across individuals, measuring expenditure in METs or calories requires methods such as doubly labeled water or indirect calorimetry, which can measure energy expenditure (Dale, Welk, and Matthews). While precise, these methods require either respiratory gas measurement (in the case of indirect calorimetry) or having subjects ingest stable isotopes of water (in the case of double labeled water). To link physical activity to the built environment, researchers will typically need survey data that links large numbers of persons to a range of built environments. Thus techniques that require medical monitoring are likely to be too costly and inconvenient. For that reason, a common technique is to calibrate survey-based measures of physical activity against laboratory measures of activity. The idea is to ask persons how often they engage in various types of physical activity, and for how long, and then use that information to measure individual physical activity.

Hence the choices for physical activity measure lie along a spectrum, with direct measurement of energy expenditure on one end, and self-reported survey questions that ideally can be mapped to direct energy measurements on another end. In between, there are techniques that provide objective measurements of activity without relying on self-reports, even if full energy expenditure is not directly measured. These include pedometers, which count steps and so can measure walking distance without relying on a survey subject's recall, and accelerometers, which can measure a broader range of movement than pedometers and so can count not merely steps but intensity and some forms of exercise (Dale, Welk, and Matthews, 2002).

Because the direct measurement of energy is precise but expensive and cumbersome, while self-reports are relatively inexpensive and convenient, large-scale studies typically rely on some form of self-reported data.[7] In some cases, self-reported data are calibrated against more direct measurement of energy expenditure, to assess the validity of the survey questions – whether or not surveys measure activity in an accurate fashion. Reliability of the survey questions is sometimes also examined – whether the survey questions yield the same response in different settings if persons are engaging in the same activities. For a discussion, see, e.g., Mahar and Rowe (2002), Matthews (2002), or Kriska and Casperson (1997). In other instances, researchers distribute pedometers or accelerometers to provide measurements that do not rely on self-reports. Here we focus on methods that lead to data that will be broadly available, accessible to persons from a large number of scholarly fields, and that span a broad range of geographies and hence built environments. Those characteristics favor survey data, and that is the focus of our discussion.

A feasible "gold standard" in secondary data for physical activity and built environment research would include the following characteristics:

---

[7] Another option is to use observers to record physical activity. For research that requires observing subjects in a large number of different built environments, this could be cumbersome, and so here we focus on survey techniques.

- measurement of the time, duration, and intensity of physical activity, using validated survey techniques, so that a rather complete measurement of activity as it relates to energy expenditure can be constructed
- data that are widely available to researchers, ideally maintained by an agency with resources to make the data broadly available
- national or at least broadly regional in scope, so that the survey respondents represent a large number of built environments
- data that allows geocoding of an individual survey respondent's home location, with the ability to match back from the geocode to the survey data so that responses can be linked to GIS or other locational data
- longitudinal data, ideally tracked for the same respondents over time
- ability to add questions

The above characteristics are a feasible ideal. It would be even more ideal, for example, to have data on physical activity, health outcomes, and the built environment in one data set, but such combinations are still rare, and so linking multiple data sets is likely to be the norm during the early stages of research on this topic. While the above list enumerates several ideals, data that do not have all of the above characteristics can still be useful. Many of the more promising longitudinal studies are still young, and cross-sectional surveys can be valuable. Similarly, while geocoding to the level of a street address or map coordinates would be ideal (and is often possible), more coarse geographic detail can also be useful.

Table 1, below, lists studies that, when judged by the above criteria, are the most promising sources of physical activity and health data that can be linked to measures of the built environment. The studies in Table 1 are drawn from the larger set of data sources listed in the appendix.

All of the studies in Table 1 provide somewhat comprehensive measurements of physical activity, and all are major national or regional surveys that are a source of secondary data available to the research community. For all of the data sets listed below, a researcher should be able to construct physical activity measures that include information about the frequency, duration, and intensity of the activity, making an approximation to energy expenditure possible. Thus the studies listed in Table 1 are among the best available data sources for physical activity when combining criteria of (1) complete measures of physical activity, (2) broad national or regional coverage, and (3) data that are easily accessible to the research community.[8] Note that the studies in Table 1 used different survey questions to measure physical activity, and researchers should consult each study for information about the physical activity questions or measurement methods. The appendix includes more information about the studies listed below and other data sources.

Table 1: Major Physical Activity and Health Data Sources

---

[8] Promising studies that have not yet been funded are excluded from Table 1. One example is the National Children's Study, with a target sample size of more than 100,000 children. That study is described in the appendix.

Study or Survey	National or Regional Scope	Geocoding Prospects or Geographic Detail	Longitudinal	Ability to Add Questions
NHIS	National	Currently MSA detail only	Annual cross-sections check	Yes
CHIS	California	Zip Code and nearest cross-street at secure UCLA site	Bi-annual cross-sections	check
NHANES	National	Address not currently available	Three waves, separate cross-sections [b]	No
BRFSS	National and for each state	Access varies by state [a]	Monthly cross-sections	Yes (at state or CDC discretion)
YRBSS	National	State or Region only	Bi-annual cross-sections	Check
Aerobics Center	Cooper clinic patients – broad coverage but non-random sample	Zip code and address, can request access with conditions to preserve subject anonymity	Longitudinal followup	Check
NHEFS	National	Check	Followup of individuals in NHANES I	N/A (in progress)
AddHealth	National	Address may be available at secure site	Longitudinal panel	check

Survey names are: National Health Interview Survey (NHIS), California Health Interview Survey (CHIS), National Health and Nutrition Examination Survey (NHANES), Behavioral Risk Factor Surveillance System (BRFSS), Youth Risk Behavior Surveillance System (YRBSS), Aerobics Center Longitudinal Study (Cooper Clinic Patients, sponsored by U.S. Public Health Service), NHANES I Epidemiologic Follow-Up Study (NHEFS), National Longitudinal Study of Adolescent Health (AddHealth)

[a] As an example of variation across states in geographic access to BRFSS data, California currently releases BRFSS data with identifies that are not more precise that county of residence. Oregon has examining whether zip code information can be released.

[b] Data collection for a fourth wave of NHANES is in progress.

None of the above studies include detailed information about the built environment, as all were developed before the question of links between the built environment and physical activity was commonly discussed in the research community. For that reason, all of the above studies will need to be linked to geographically specific information (GIS) about the built environment. The most efficient method will be to geocode physical activity or health survey data using a geographic information system. For that reason, the prospects for geocoding the data or, if precise geocoding information is not released, the geographic level of detail that is released to researchers is shown in Table 1.

One of two strategies will be necessary to preserve subject anonymity. Geocoding information (typically a street address) can be released in a secure facility. Researchers could load external data, including GIS data or information about the built environment from observational studies, onto the secure computers in the facility. The U.S. Census

Bureau has allowed access to sensitive data at secure sites for years, and the methods used at those sites could be a model for secure access to geocoding information in health and physical activity data. The Health Information Privacy Act (HIPA) will require that subject anonymity be preserved, but common research norms and federal funding procedures already require strong assurances of subject anonymity. Thus compliance with HIPA creates additional legal requirements, but it is not clear that HIPA would add greater complexity to secure data sites than already exists.

An alternative to secure data sites is to collect information about the built environment near each survey respondent's residence for one of the data sources in Table 1 (or a similar data source), link that information to individual health and physical activity data, and then release the information to researchers with all address and other identifying information stripped from the data released to the research community. This avoids the cost of operating a secure data analysis facility, but adds a likely larger cost in collecting and linking built environment data at agency, rather than researcher, initiative.

Given that research into the built environment and physical activity is in its early stages, allow researcher innovation in measurements of the built environment should be preferred. Thus a high funding priority is the creation and maintenance of secure data analysis sites that allows researchers access to geocoded information about individual survey respondents with the ability to link built environment data to health and physical activity survey information is a high priority. One strategy that might produce large benefits in data availability would be to select a small number of studies from the data sources listed in Table 1 and fund programs designed to increase geocoded information, allowing linking to data on the built environment, at secure sites that create access to a broad research community.

IV. Built Environment Data

Data on the built environment can be grouped into four categories, shown below.

- ♦ Standardized land use data collected by agencies and made available to the research community
- ♦ Land use or built environment data constructed by researchers, using GIS or remote sensing technology
- ♦ Observational data, collected by walking or driving neighborhoods and recording characteristics of the built environment
- ♦ Self-reported data about the built environment from survey respondents

Standardized land use or built environment data that can be used in an off-the-shelf fashion are still rare. One example of such data is Portland's Regional Land Information System (RLIS), which is described later in this section. Yet few metropolitan areas measure land use and the built environment in ways that combine detail and convenience as RLIS does, and built environment measures are not standardized across different data sources. For those reasons, researchers will typically need to combine physical activity data with specially constructed measures of the built environment. Lacking broadly available, standardized built environment data, this section focuses on how the research community can construct measures of the built environment, and how those measures might be standardized at some future time.

For that reason, we focus on constructing built environment data either through automated resources (GIS or remote sensing) or by observing and recording characteristics. Self-reported data available in surveys will often be less useful, since self-reports do not provide objective measures of the built environment. Also, individuals are a poor judge of some built environment characteristics, such as density, implying potential reliability and validity problems with self-reported measures. Self-reported measures of the built environment might play some role, especially once survey techniques in that area have been calibrated against objective measures to understand validity and reliability, but for now we suggest that objective measurement, either through automated techniques or by direct observation, will be the most commonly used tool for measuring the built environment.

The distinction between automated and observational data highlights a fundamental tension in measuring the built environment. If research is based on observations on a large number of individuals who live in many locations, automated data collection is often the only feasible option, since have a research team directly observe possibly thousands or tens of thousands of neighborhoods would be costly.[9] Direct observation of the built environment, often called an "audit" in the public health literature, is typically used when the units of observation are neighborhoods, and the number of neighborhoods is chosen to be small and consistent with the effort required to observe and code built environment features in study areas. On the other hand, automated measures of the built environment typically limit the researcher to what is available via GIS or, in rare cases, remote sensing, while observational methods can be tailored for more rich detail. The tradeoff then is, to some extent, detail at a cost that often precludes studying individuals dispersed across many neighborhoods. Because the literature should favor studies of individuals, for purposes of furthering our understanding of human behavior and causal links, automated data collection for built environment characteristics will likely be favored in many studies. Yet below we describe both data collection methods, because direct observation will play a role possibly in validating automated data collection or in supplementing automated data for small samples.

A. Automated Built Environment Data

Studies of the built environment can borrow techniques to measure land use that were developed within the context of land use – travel behavior studies. Linking individual survey data to GIS measures of the built environment has become common in land use – travel behavior studies (e.g. Cervero and Kockelman, 1997; Boarnet and Sarmiento, 1998; Crane and Crepeau, 1998; Boarnet and Crane, 2001). Earlier studies in the land use – travel behavior literature predated GIS technology, but used similar methods for measuring land use by hand (e.g. Vickerman, 1972; Hanson and Hanson, 1981).

---

[9] Note that the sample sizes for some of the physical activity data sources listed in the appendix range from tens of thousands to over a hundred thousand individuals.

Cervero and Kockelman (1997) first popularized the concept of the "three D's" of built environment measurement – density, diversity, and design – and most later studies in land use – travel behavior adopted techniques that are compatible with that typology. Density typically includes both population and employment densities, diversity includes measures of land use mix, in particular the mix of commercial and residential land uses within a neighborhood, and design is most commonly regarded as the street network. In terms of street networks, measures of street grids, which can be operationalized based on measures of four-way intersections as a percentage of total intersections within an area (e.g. Kulkarni and McNally, year; Boarnet, Nesamani, and Smith, 2003) are most common.

Note that developing these measures requires linking to multiple geographic data sources. This includes census data (e.g. for population or housing densities or characteristics), local land use or employment data (for measures of diversity), and measures generated by GIS base maps for street design. Employment and land use data are typically available from state or local authorities, and obtaining those on a national level will be a challenge. Zip code employment data are available from the Census Bureau's County Business Patterns, but land use data are typically collected only by local governments or metropolitan planning agencies. For that reason, recent research has often turned to the U.S. Government's National Resources Inventory, which tracks land use changes and can be used to measure the amount of urbanized land within a metropolitan area. More fine-grained distinctions based on land use within parcels typically requires local data.

The built environment measures developed to study travel behavior will need to be adapted to studies of physical activity. Until recently, studies of land use – travel behavior focused on motorized travel (automobiles and transit).[10] Features such as sidewalks, parks, public spaces, crime rates, and even street lighting might influence physical activity but likely be less important determinants of driving or transit travel behavior. So some care must be taken to adopt the techniques from land use – travel behavior studies in ways that measure elements of the built environment that likely influence physical activity.

Notice that a focus on physical activity will likely require additional geographic data sources that were not typically used to measure the built environment in travel behavior studies. Crime rates can be obtained from local law enforcement agencies or the state agency charged with collecting Uniform Crime Report data. As with other data, these should be at a level of geographic detail that is small enough to capture variation across locations – census tracts have been commonly used as the level of geographic detail in the past.

Infrastructure that supports physical activity – sidewalks, public spaces, parks, school locations, or open spaces – can often be obtained from GIS maps that are either commercially available or maintained by local governments. As was mentioned above,

[10] Recent exceptions that examine land use and non-motorized travel behavior include Clifton and Handy, 2000 and Greenwald and Boarnet, 2001.

such data sources are not standardized across cities or metropolitan areas, and will typically have to be constructed by the researcher. In some localities, needed data may not be available if, for example, public spaces or sidewalk infrastructure are not catalogued and available in GIS maps. Below we describe two possible emerging solutions to this problem.

The first solution is the development of GIS-based land and built environment databases. One of the better examples of such systems is the Regional Land Information System (RLIS) for metropolitan Portland, Oregon. The Portland Metro has been developing this sophisticated GIS land use database for over ten years. The RLIS database includes GIS data on sidewalks, bike routes, rivers, paths, vegetation cover, slope, parks, and open space, linked to both street and census geography. This is an advanced set of geographic data, and allows researchers to use measures of the built and natural environments without having to develop those measures on their own.[11] The RLIS database is likely a prototype of advanced GIS land use data banks that will be developed in other urban areas. For now, few systems can match the broad range of data that are easily available to the research community in RLIS. As other cities and metropolitan areas develop similar land use databases, the labor needed to measure the built environment will drop. What now often requires custom-designed built environment data constructed by the researcher might someday be available from data banks like RLIS. Having said that, there is currently no standardized method of measuring or cataloging these variables, and no centralized national repository of such data, so that even when other communities implement systems like RLIS, researchers may need to canvass GIS-based built environment data from several urban areas.

The second solution is the use of remote sensing or aerial photography data to measure features that might not be present in GIS data banks for particular regions. Features such as sidewalk coverage, parks, or vegetation cover might not be available in GIS format in some metropolitan areas. In fact, the availability of such data depends in part on whether local agencies have collected and geocoded the information. In cases where GIS information is not available, the alternatives are either to directly observe the environment (as in audit studies discussed below) or to use remote sensing or aerial photography data that can provide measurements of the built environment.

As part of the National Spatial Data Infrastructure initiative, a Geospatial One-Stop Portal has been created and can be accessed at www.geodata.gov. This portal is a convenient starting point from which to access nationwide geospatial data from a multitude of available sources, including satellite imagery and aerial photography. In most cases, data can be downloaded from the portal as GIS files that can be viewed and used with GIS software. The portal also provides metadata (data about the contents of the files, file attributes, and contact information), and has a map viewer that enables viewing of available maps online.

As a sample case, to find geospatial characteristics of a neighborhood in a city, the simplest approach is to run a general search in www.geodata.gov using the city's name,

[11] For more about RLIS, see Bolen (2002).

e.g., “Los Angeles”, as keywords, while using the most generic search parameters (i.e., using “any” for all the other choices). This search would generate a list of available data – some of which could be automatically viewed or downloaded while others would have to be obtained directly from the source (provided in the details and/or metadata section of the search results). From the data provided, one can determine whether the data are relevant. To use the data, a researcher can either download the files in GIS format if available online or contact the source for access to the data.

B. Built Environment Data from Direct Observation

In cases where GIS data are not available or do not capture relevant features of the built environment, researchers can directly observe the environment, often by walking neighborhoods and recording information about selected characteristics using an environmental audit instrument. Likely the first environmental audit instrument developed to measure built environment features that are associated with physical activity is the Systematic Pedestrian and Cycling Environmental Scan Instrument, or SPACES (Pikora et al., 2002).

To use the SPACES audit tool, observers walk neighborhoods answering questions that prompt them to record information about street width, sidewalks, traffic volume, lighting, aesthetics, parks and shops, and various other factors that might be linked to physical activity. Information is recorded for individual blocks, and so can be aggregated to higher geographies or analyzed at the block level. The SPACES audit tool has been reliability tested, and Pikora et al. (2002) report that many of the questions have high inter-rater reliability. A similar audit tool, also applied at the block level, was developed to measure the built environment near school sites in the evaluation of the California Safe Routes to School program (California Safe Routes to Schools, December, 2003).

The advantage of an audit or direct observation tool is that it gives researchers an opportunity to measure characteristics that might not be available in GIS, remote sensing, or aerial photography data. Yet environmental audits require additional labor to observe each built environment, and can yield an amount of data that might be unwieldy in certain circumstances. The tradeoff in data management and labor cost across observational and automated built environment data likely will vary in part with problem context and in part based on the resources and skills available to a research team. A team adept at GIS data might view an environmental audit as unduly time consuming, while a team with access to research assistants but with few GIS skills might find the audit quicker and cheaper than relying on GIS data. As was mentioned above, linking to a large number of individual observations in a broad range of locations likely will favor automated data collection, but environmental audits can be used to supplement GIS data for smaller samples and to compare with GIS measures of the built environment.

V. Land Use – Travel Behavior Data and Physical Activity Research

[This section to be completed]

References [To be completed]

Boarnet, M. and Greenwald, G. 2000. Land Use, Urban Design, and Non-Work Travel: Reproducing for Portland, Oregon Empirical Tests from Other Urban Areas. *Transportation Research Record* 1722: 27-37.

Boarnet and Crane. 2001a. The Influence of Land Use on Travel Behavior: Specification and Estimation Strategies.

Boarnet, M. and Crane, R. 2001b. *Travel by design: The influence of urban form on travel.* New York: Oxford University Press.

Boarnet, M. and Greenwald, M. 2000. Land Use, Urban Design, and Non-Work Travel: Reproducing for Portland, Oregon Empirical Tests from Other Urban Areas. *Transportation Research Record*, number 1722, pp. 27-37.

Boarnet, Marlon G. and Sharon Sarmiento. 1998. Can Land Use Policy Really Affect Travel Behaviour? A Study of the Link Between Non-Work Travel and Land Use Characteristics. ***Urban Studies*** 35, 7: 1155-1169.

Cervero, R. and Kockelman, K. 1997. Travel demand and the 3Ds: density, diversity and design. Transportation Research D, 2(3):199-219.

Greenwald and Boarnet. 2001. Built Environment as Determinant of Walking Behavior: Analyzing Nonwork Pedestrian Travel in Portland, Oregon. *Transportation Research Record* number 1780, pp. 33-42.

Handy, Susan L., Marlon G. Boarnet, Reid Ewing, and Richard E. Killingsworth. 2002. How the Built Environment Influences Physical Activity: Views from Urban Planning. *American Journal of Preventive Medicine* 23,2: 64-73.

Handy, Susan and Kelly Clifton. 2001. Local Shopping as a Strategy for Reducing Automobile Travel. *Transportation* 28,4: 317-346.

Handy, Susan. 1993. Regional Versus Local Accessibility: Implications for Nonwork Travel. *Transportation Research Record*, number 1400: 58-66.

Hanson, Susan and Perry Hanson. 1981. The Travel-Activity Patterns of Urban Residents: Dimensions and Relationships to Sociodemographic Chateristicis. *Economic Geography* 57,4: 332-347.

Krizek, Kevin. 2003. Residential Relocation And Changes In Urban Travel: Does Neighborhood-Scale Urban Form Matter? *Journal of the American Planning Association* 69,3. (forthcoming)

Vickerman, Roger W. 1972. The Demand for Non-Work Travel. *Journal of Transport Economics and Policy* 6,2: 176-210.

Bolen, Richard. 2002. GIS: Essential for Urban Growth Management: Portland, Oregon Metropolitan Area. Presented at RLIS @ Ten Symposium, Portland, Oregon, March. <http://www.metro-region.org/library_docs/maps_data/gis_and_planning.pdf>, accessed November 18, 2003.

Paffenbarger, R.S., Hyde, R.T., Wing, A.L., & Hsieh, C.C. (1986). Physical activity, all-cause mortality, and longevity of college alumni. *New England Journal of Medicine, 314*, 605-613.

Pate, R., Pratt, M., Blair, S., Haskell, W., Macera, C., Bouchard, C., Buchner, D., Ettinger, W., Heath, G., King, A., Kriska, A., Leon, A., Marcus, B., Morris, J., Paffenbarger, R., Patrick, K., Pollock, M., Rippe, J., Sallis, J., and Wilmore, J. (1995). Physical activity and public health: a recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *Journal of the American Medical Association*, 273(5): 402-407.

Leon, A.S., Connett, J., Jacobs, D.R., & Rauramaa, R. (1987). Leisure-time physical activity levels and risk of coronary heart disease and death: the Multiple Risk Factor Intervention Trial. *Journal of the American Medical Association, 258*(17), 2388-2395.

Ekelund, L.G., Haskell, W.L. Johnson, J.L., Whaley, F.S., Criqui, M.H., & Sheps, D.S. (1988). Physical fitness as a predictor of cardiovascular mortality in asymptomatic North American men: the Lipid Research Clinics mortality follow-up study. *New England Journal of Medicine, 319*, 1379-1384.

Blair, S.N., Kohl, H.W., Paffenbarger R.S., Clark D.G., Cooper K.H., & Gibbons L.W. 1989. Physical fitness and all-cause mortality: a prospective study of healthy men and women. *Journal of the American Medical Association, 262*(17), 2395-2401.

Morris, J.N., Clayton, D.G., Everitt, M.G., Semmence, A.M., & Burgess, E.H. (1990). Exercise in leisure time: coronary attack and death rates. *British Heart Journal, 63,* 325-334.

Sandvik, L., Erikssen, J., Thaulow, E., Erikssen, G., Mundal, R., & Rodahl, K. (1993). Physical fitness as a predictor of mortality among healthy, middle-aged Norwegian men. *New England Journal of Medicine, 328,* 533-537.

U.S. Department of Health and Human Services. *Physical Activity and Health: A Report of the Surgeon General.* Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, 1996.

Appendix: Data Sources

Physical Activity and Health Data Sources: Cross-Sectional or Stacked Cross-Sections

1. Behavioral Risk Factor Surveillance System (BRFSS) (www.cdc.gov/brfss/about.htm)

Brief Description: The BRFSS was established by CDC in 1984 to monitor state-level personal health behavior. States use the CDC-developed questionnaire and may add their own questions. The survey collects data on behaviors that are linked with the leading causes of death in the US – heart disease, stroke, cancer, diabetes. Therefore, data on physical activity, diet, tobacco and alcohol use, and use of preventive care are recorded.
Sponsor Agency: CDC
Sample Size:
Data collection method: monthly phone interviews
Study Area: each of the 50 states, plus District of Columbia and three territories
Survey interval: monthly (15 states since 1984; all 50 states since 1994)
Data obtained:
Frequency and duration of leisure-time physical activity, occupational activity, measure of perceived safety from crime
Access to Geographic Data
Varies by state. In California, county level data only according to Holly Hoegh. In Oregon, zip code data exist but are currently not released. However, according to Kathryn Pickle, they are discussing the feasibility of releasing the data, so there is hope! One possible issue is the implication of the Health Information Protection Act.
Contact Persons: contact persons in each state

2. Youth Risk Behavior Surveillance System (YRBSS) (www.cdc.gov/nccdphp/dash/yrbs/index.htm)

Brief Description: The YRBSS was established by CDC in 1990 to monitor youth health risk behavior. National, state, and local school-based surveys, which constitute the surveillance *system* collect data on behaviors that are linked with the leading causes of death, disability, and social problems in the US. Therefore, data on physical activity, diet, tobacco and alcohol use, and use of preventive care are recorded.
Sponsor Agency: CDC
Contact: Joanne Grunbaum, jpg9@cdc.gov, 770-48-6182
Sample Size: average per regional survey is approximately 1800 persons representative of study area; the Youth Risk Behavior Survey administered in 1992 obtained data for 11000 persons ages 12 to 21.
Data collection method: in-school survey administration
Study Areas: in 2001, 38 states, 19 metropolitan areas, and 7 territories participated; national, state, and local surveys are representative of 9th through 12th grade students in public and private high schools in the study area.
Survey interval: every two years
Data obtained:

- Physical activity: number of days per week with at least 20 minutes of vigorous exercise, at least 30 minutes of moderate exercise, participation in sports teams, PE class participation, and muscle-strengthening.exercises.
- Health conditions: covers major types of diseases, vital statistics, other risky behavior besides physical inactivity (diet, alcohol, cigarettes)

Access to Geographic Data

None except for state of region according to Joanne Grunbaum.

Other comments:

- As part of YRBSS, the National College Health Risk Behavior Survey (NCHRBS) was conducted in 1995 with sample size of 4609 undergraduate students 18 years and older, representative of US public and private universities and colleges.
- Also part of YRBSS was the National Alternative High School Youth Risk Behavior Survey administered to 9000 students nationwide, representative of students in alternative high schools.

3. National Health Interview Survey (NHIS)

Sponsor Agency: National Center for Health Statistics, NCHS, CDC

Contact Persons: Viola Lang, 301-458-4901

Sample Size (2001): 38,932 households, 100,761 persons, 39,633 families

Collected by: US Census Bureau

Data collection method: Personal household interview

Study Area: USA

Survey interval: annual

Data obtained:

- Physical activity: frequency, duration, type, intensity
- Health conditions: covers major types of diseases, vital statistics, other risky behavior besides physical inactivity (diet, alcohol, cigarettes)

Access to Geographic Data

Only MSA's according to Viola Lang.

Other Comments:

- Supplementary questions should be very easy to add. The last re-design of the survey was done in 1997.

4. National Health and Nutrition Examination Survey (NHANES) (http://www.cdc.gov/nchs/nhanes.htm)

Sponsor Agency: National Center for Health Statistics, NCHS, CDC

Contact: Ken W. Harris, Research Data Center, kwh1@cdc.gov, 301-458-4277

Sample Size: Most recent dataset 1999-2000: 9965 persons; NHANESIII – 1988-1994: 14,827 adults at least 20 years old responded to the 211 behavioral and demographic variables but 40,000 persons in dataset; NHANES I and II had 28,000 persons in the sample, ages 1 – 74 in the former, and ages 6 months-74 in the latter.

Collected by: National Center for Health Statistics (NCHS)

Data collection method: home interview and health tests in mobile lab

Study Area: USA

Survey interval: Current NHANES is from 1999-2003; prior surveys: NHANES III from 1988-1994, NHANES II from 1976-1980, NHANES I from 1971-1975, and Hispanic HANES (HHANES) from 1982-1984
Data obtained:

- Physical activity: frequency, duration, type (includes whether walked or biked to work; walking or biking also covered under exercise), intensity
- Health conditions: covers major types of diseases, vital statistics, other risky behavior besides physical inactivity (diet, alcohol, cigarettes)
- Locale: home characteristics (type, age, size, paint condition, tap water treatment, etc.), rural/urban/suburban, region in USA

Comments:

- No geographic information except home address; if walking or biking to work, have employer's name but not address.
- Ongoing discussions for expanding survey and increasing its usability include: longitudinal HANES, types of community indicators and ecological variables, and ways to link data with other data sources.

5. California Health Interview Survey (CHIS) (www.chis.ucla.edu)
Sponsor Agency: California State Department of Health Services
Contact: Lee Habte, CHIS Data Access Center, lhabte@ucla.edu, 310-794-2684
Sample size: 55,248 adults, 5801 teens, 12,592 children
Survey interval: 2001 was first; planned for every 2 years (2003 survey to start next month)
Data obtained: physical activity (how much walking/biking, daily level of physical activity, leisure-time level of physical activity, muscle strengthening activity)
Access to Geographic Data
Zip code for all data and nearest cross-streets for San Diego and Los Angeles data – available by application to Data Access Center.
Other Comments

- Year 2003 survey will have geocoding.
- Lots of helpful information to researchers (click on "Researchers" button) on website.

6. US Physical Activity Study
Sponsor Agency: CDC
Contact Person: Dr. Ross C. Brownson, brownson@slu.edu
Sample Size: 1818 adults
Collected by: St. Louis University
Data Collection Method: Phone interview
Study Area / Population: Nationwide representative sample of adults at or above 18 years age, oversampled low-income individuals
Survey date: interviews completed between September 1999 and January 2000
Relevant data obtained: general access to places of exercise; access to specific types of recreational activity, neighborhood characteristics (e.g., existence of sidewalks, hills, safety, lighting); attitudes towards exercise and policy related to physical activity; perception of amount of physical activity in neighborhood

Access to Data: principal investigator would need to request permission from Dr. Brownson; have telephone numbers that could be matched to street addresses.

Physical Activity and Health Data Sources: Longitudinal

1. NHANES I Epidemiologic Follow-Up Study (NHEFS) (www.cdc.gov/nchs/about/major/nhefs/nhefs.htm)
Sponsor Agency: National Center for Health Statistics, National Institute of Aging, other components of National Institute of Health, Substance Abuse and Mental Health Services Administration, other centers within CDC
Contact: Paula Norman, 301-458-4530
Sample size: 14,407 adults who were 25 – 74 years at the time of NHANES I (1971-75) examination
Survey interval: 1st wave in 1982-84; 2nd and 3rd waves in 1987 and 1992 respectively for entire non-deceased cohort; special survey for those 55-74 years at baseline exam in 1986.
Relevant data obtained: number of years, frequency, and duration of regular exercise/sports, amount of light physical activity, height, weight
Access to Geographic Data
Attrition Rate

2. The National Longitudinal Study of Adolescent Health (AddHealth) (www.cpc.unc.edu/addhealth/study.html)
Sponsor Agency: National Institute of Child Health and Human Development plus other national health institutes and offices of Department of Health and Human Services
Contact: Joyce Tabor, Data Manager, 650-966-4487
Sample size: 90,118 students (Wave I in-school surveys), 20,745 adolescents (Wave I in-home surveys)
Study population: students from grades 7 – 12 in 145 middle/junior high/high schools as a representative of US schools in terms of region, degree of urbanicity, school size, ethnicity, school enterprise (public/private/religious)
Survey interval: 1st wave in 1994-1995; 2nd wave in 1996; 3rd wave in 2001-2002
Relevant data obtained: from students – health status and health-related behaviors including physical activity (frequency of: exercise, participation in active sport, biking/rollerblading or skating/skateboarding), from parents – attitude towards and perception of neighborhood and school
Access to Geographic Data: all home locations geocoded and linked to block group census areas (either have addresses, zip +4, or GPS) not released to researchers, but researchers might be able to request to work with the data within the security vault.
Response Rate: 88% in Wave II, 77% in Wave I
Other Comments: Oversampled Chinese, Cuban, Puerto Rican, and disabled students as well as African-American students from well-educated families.

3. Nurses' Health Study (www.channing.harvard.edu/nhs/)
Sponsor Agency: National health institutes

Contact: Gary Chase, gary.chase@channing.harvard.edu, 617-525-2279
Sample size: 122,000 women in NHS1, 116,700 in NHS2
Study population: In NHS1, female married registered nurses ages 30-55 in 1976 in 11 of the most populous states that released registered nurse contact information (CA, CT, FL, MD, MA, MI, NJ, OH, PA, TX); in NHS2, female registered nurses ages 25-42 in 1989 again from the most populous states that released contact information (CA, CT, IN, IA, KY, MA, MI, MO, NY, NC, OH, PA, SC, TX)
Survey interval: every two years (since 1976 for NHS1, since 1989 for NHS2), but physical activity questions asked every four years
Relevant data obtained: health status and health-related behaviors including physical activity (walking pace, flights of stairs, recreational activity frequency and duration)
Access to Geographic Data: zip code data available; also have addresses but not currently in datasets, so would have to request that data be entered into database.
Attrition Rate: have maintained about 90% response rate for both NHS1 and NHS2

4. The Health Professionals Follow-Up Study (www.hsph.harvard.edu/hpfs)
Sponsor Agency: National Heart, Lung, & Blood Institute and National Cancer Institute
Contact: Dr. Walter Willett, walter.willett@channing.harvard.edu
Sample size: 51,529 men
Study population: about 30,000 dentists, 4000 pharmacists, 4000 optometrists, 2000 osteopath physicians, 1600 podiatrists, 10,000 veterinarians
Survey interval: every two years since 1986
Relevant data obtained: health status and health-related behaviors including physical activity (walking pace, flights of stairs, recreational activity frequency and duration)
Access to Geographic Data: residences exist by geocodes but cannot allow identification of households because of confidentiality restrictions; apply to Dr. Willett for data access.

5. The College Alumni Health Study (www.stanford.edu/~paff/CAHSPrecis.html)
Contact: Dr. I-Min Lee, i-min.lee@channing.harvard.edu, 617-278-0806
Sample size: over 60,000 in 1960
Study population: University of Pennsylvania and Harvard College alumni who were students between 1916 and 1950 (born between 1896 and 1934)
Survey interval: approximately every 5 years since 1960
Relevant data obtained: health status and health-related behaviors including physical activity (walking – frequency, duration, intensity; stair-climbing; sport/recreation/other physical activity; daily activity breakdown from most vigorous to sleeping/reclining activities)
Access to Data: principal investigator would need to request permission from Harvard School of Public Health for access to any data; have zip code and address data but current research already underway on relationship between urban form, physical activity, and health outcomes so such data are not releasable at this time.

6. The Aerobics Center Longitudinal Study (http://www.cooperinst.org/respub.asp)
Sponsor Agency: US Public Health Service
Contact: Beth Wright, bwright@cooperinst.org, 972-341-3246
Sample size: 75,000 individuals

Study population: Cooper Clinic patients
Survey interval: about every 4 years since 1970
Relevant data obtained: frequency, duration, and intensity of walking/running/swimming/biking activities
Access to Geographic Data: zip codes and addresses available and data release to researchers possible upon approval
Other Comments: research on topic of relationship between urban form, physical activity, and health outcomes currently underway

7. Baltimore Longitudinal Study of Aging (www.grc.nia.nih.gov/branches/blsa/blsa.htm)
Sponsor Agency: National Institute on Aging (NIA)
Contact person: Robin Campbell, campbellro@grc.nia.nih.gov; 410-350-3950
Sample size: 3100
Study population: consists of volunteers who join and quit as desired; participants are at least 20 years old
Survey interval: started in 1958; depends on age, at age 80 and above, participant is asked every year
Relevant data obtained: estimate of amount of time spend doing each of 100 activities
Access to Geographic Data: most likely not

8. National Children's Study (www.nationalchildrensstudy.gov)
Contact person: Dr. Ross Brownson, brownson@slu.edu, 301-594-9147
Sponsor agency: not yet funded
Target sample size: >100,000 children
Survey interval: not yet defined, but intent is to follow children from birth until age 21
Study goal: examine effects of environmental influences on children's health and development, including effects of physical surroundings, geographic locations, and behavioral influences and outcomes

Built Environment Data Sources

RLIS: Portland Metro Regional Land Information System
a geographical database that covers 24 cities within the 3 counties of Portland Metro
GIS spatial data include among many:
Boundary lines between cities, counties, neighborhoods
Census tracts
Sidewalks, bike routes, rivers, multi-use paths, transit routes
Major contours, slope, vegetation cover, parks and open space
Contact: Data Resource Center at (503) 797-1742, drc@metro-region.org or see descriptions of RLIS at the Portland Metro web site, www.metro-region.org