

Are Your Asset Data as Good as You Think? Conducting a Comprehensive Census of Built Assets to Improve Urban Population Health

Jennifer A. Makelarski, Stacy Tessler Lindau,
Vanessa D. Fabbre, Colleen M. Grogan, Eugene M. Sadhu,
Jonathan C. Silverstein, Thuy Tien T. Tran,
Martha Van Haitsma, Eric Whitaker, and Daniel Johnson

ABSTRACT *Secondary data sources are widely used to measure the built asset environment, although their validity for this purpose is not well-established. Using community-engaged research methodology, this study conducted a census of public-facing, built assets via direct observation and then tested the performance of these data against widely used secondary datasets. After engaging community organizations, a community education campaign was implemented. Using web-enabled cell phones and a web-based application prepopulated with the secondary data, census workers verified, modified, and/or added assets using street-level observation, supplementing data with web searches and telephone calls. Data were uploaded to <http://www.SouthSideHealth.org>. Using direct observation as the criterion standard, the sensitivity of secondary datasets was calculated. Of 5,773 assets on the prepopulated list, direct observation of public-facing assets verified 1,612 as operating; another 653 operating assets were newly identified. Sensitivity of the commercial list for nonresidential, operating assets was 61 %. Using the asset census as the criterion standard, secondary datasets were incomplete and inaccurate. Comprehensive, accurate built asset data are needed to advance urban health research, inform policy, and improve individuals' access to assets.*

KEYWORDS *Urban health, Built environment, Asset mapping, Asset census, Community-engaged research methodology, Population health*

Health inequities across socioeconomic and racial/ethnic strata in the USA are largely explained by social and built environmental factors that extend beyond the health care system.¹⁻⁶ Recent studies have shown important variation in health status and health-related behaviors by the differences in the built asset environment,⁷⁻¹³ such as density

Makelarski and Lindau are with the Department of Obstetrics and Gynecology, The University of Chicago, Chicago, IL, USA; Lindau is with the The MacLean Center for Clinical Medical Ethics, Department of Medicine—Geriatrics, The University of Chicago, Chicago, IL, USA; Fabbre and Grogan are with the Center for Health Administration Studies, School of Social Service Administration, The University of Chicago, Chicago, IL, USA; Sadhu is with the Department of Surgery, The University of Chicago, Chicago, IL, USA; Sadhu and Silverstein are with the Computation Institute, The University of Chicago, Chicago, IL, USA; Makelarski, Lindau, Tran, Whitaker, and Johnson are with the Urban Health Initiative, The University of Chicago, Chicago, IL, USA; Haitsma is with the Survey Lab, The University of Chicago, Chicago, IL, USA; Whitaker is with the Department of Medicine, The University of Chicago Medical Center, Chicago, IL, USA; Johnson is with the Department of Pediatrics, The University of Chicago Medical Center, Chicago, IL, USA.

Correspondence: Jennifer A. Makelarski, Department of Obstetrics and Gynecology, The University of Chicago, Chicago, IL, USA. (E-mail: jmakelarski@babies.bsd.uchicago.edu)

of fast food restaurants. Using data purchased from Dun and Bradstreet,¹⁴ Boone-Heinonen et al. found a positive relationship between fast food restaurant density and self-reported fast food consumption, but no relationship between supermarket density and self-reported fruit and vegetable intake.⁸ Block and colleagues supplemented data from Dun and Bradstreet¹⁴ with telephone book entries and government inspection reports to longitudinally assess the impact of proximity to fast food restaurants and grocers on body mass index (BMI) in the Framingham Heart Study Offspring Cohort.⁷ They found that residential proximity to fast food restaurants and grocers was significantly associated with BMI only in women; each 1 km increase in the driving distance to a fast food restaurant and to a grocer was associated with a 0.19- and 0.11-unit decrease in BMI, respectively.

To date, studies like these have almost exclusively relied on secondary data sources (e.g., market research lists, government reports)^{7-13,15-21} to measure the built asset environment. Secondary data sources are, in most regions, the only asset data readily available at a low cost. Studies typically acknowledge the potential limitations of secondary data sources,^{7,8} but few studies have investigated the validity of secondary data sources.²²⁻²⁴ Whether secondary data sources are a valid measure of the built asset environment, particularly in low-income, urban communities, is not well understood and may explain inconsistent findings across studies. In the USA, Bader et al. examined the concordance of data from InfoUSA²⁵ and direct observation for the presence or absence of any built food assets and found low to moderate agreement (kappa, 0.32–0.70).²² Notably, the authors examined concordance of the presence or absence of any food asset, rather than concordance on each specific asset, and did not verify if assets identified by InfoUSA were operating. In the UK, Lake et al. found that, using direct observation as the criterion standard, the sensitivity of government secondary data sources in identifying food assets was relatively high (84 %), whereas the sensitivity of commercial secondary data sources (yellow pages and online directory) was substantially lower (53 and 51 %, respectively).²⁴ Additionally, 59 % of the built food assets in the online directory were not present in the field.

To our knowledge, no comprehensive (including all sectors) census of built assets using direct, street-level observation has been conducted nor has the validity of secondary data sources in capturing all assets among all sectors of the built asset environment been tested. Without an empirical, comprehensive measure of the built asset environment, the ability to study the impact of the intersectoral^{12,26} built environment on health is limited (e.g., Does the density of fitness centers mediate the relationship between fast food density and BMI?). An empirical measure is necessary to systematically identify the biases inherent in secondary data sources.

In addition to population health research applications, comprehensive asset data have high translational value for providers who make referrals for unmet social needs or who prescribe behavior changes that directly relate to community assets (e.g., increase exercise, eat more fruits and vegetables, see a counselor, etc.). In a recent survey by the Robert Wood Johnson Foundation, four out of five physicians surveyed were not confident in their ability to address their patients' social needs and expressed a desire to write prescriptions to fill this gap.²⁷ A census of assets may be useful in filling this gap, provided that the data are readily available to providers and are accurate.

Building on work implemented to create a coordinated system of care (<http://www.uchospitals.edu/programs/community/programs/sshc.html>), the University of Chicago's academic medical center on Chicago's South Side incepted the Urban

Health Initiative.^{27,28} Its vision is for Chicago's South Side to be a model of excellent urban health by 2025, using an asset-based approach.²⁹ The research arm of the Urban Health Initiative aims to learn how built assets in low-income urban communities can be optimized for health. In service to this aim, the research team sought to accurately and completely measure the South Side's built asset environment using direct observation. Building on prior work,^{22,24,30-32} the team developed a community-engaged asset-based methodology for a comprehensive, directly observed census of all public-facing, built assets in this dense, low-income urban area. The team aimed to create a census that served a dual purpose—creation of a list of built assets that could be used to advance knowledge about the relationship between the built asset environment and health and a map that could accurately direct individuals to the assets they need. This paper describes the asset census methodology and the performance of these data in comparison to other widely used secondary data sources of built assets.

METHODS

Study Region and Research Team

Chicago is divided into 77 distinct community areas; each includes one or more census tracts.³³ Thirty-four of these communities constitute Chicago's South Side and the primary service area of the University of Chicago's medical center. The South Side's (95 mi²) population of about 800,000 is predominantly African American and suffers from endemic unemployment (13 % unemployment rate), high poverty (50 % of households at or below 200 % of the federal poverty level), and health inequities (e.g., 28 deaths per 100,000 due to cardiovascular disease versus 21 per 100,000 overall in Chicago).^{34,35} The asset census was piloted in six (11 mi²) South Side communities. The project used data generated via publicly available observations and was deemed exempt by the university's Institutional Review Board.

The research team was composed of a multidisciplinary team of university members (e.g., physicians, social scientists, informaticians), leadership from the University of Chicago's Survey Lab contracted to carry out field operations, and the University's Computation Institute contracted to build data collection tools and the website. After 10 formative meetings, 12 community advisors (2 from each community), identified by the Medical Center's Office of Community Affairs and the Urban Health Initiative's team, were invited to join an advisory board and serve as the "community voice" in the development and implementation of study methodology. Invitees were residents, business owners, and community leaders with diverse areas of expertise (e.g., community organizing, health, economic development). Many had existing relationships with the university.

Community Education Campaign

Prior to the start of data collection, the research team instituted a community education campaign to generate publicity, foster positive interactions with community residents, establish legitimacy, and alert the community to the asset census workers in their communities. With extensive Survey Lab and community advisor input, we generated a one-page flyer and tri-fold brochure in English and Spanish that included a project description, an example asset map, research team contact information, and a picture of the census workers in uniform. Research team leaders met with Aldermanic office representatives, police, and other community leaders.

Asset census workers distributed flyers in each community 1 week prior to data collection. The team sent mass emails using distribution lists provided by the office of community affairs and community organizations. The project also received unsolicited positive local press.³⁶

Asset Eligibility

Built “assets” were defined as physical places providing goods and/or services: (1) for-profit businesses (e.g., grocers, manufacturers), (2) nonprofit organizations (e.g., health clinics, churches), (3) government entities (e.g., libraries, police stations), (4) programmed residences (e.g., nursing homes, college dorms), and (5) licensed in-home daycares. Other residential buildings (e.g., public housing, rental housing) and outdoor public spaces (e.g., parks) were excluded. Listings identified as consumers of private services that require an employee identification number, for example, elderly persons hiring nurses, families hiring nannies, or households that hire cleaning services, were also excluded. Due to resource limitations, in-field data collection was restricted to assets with a public presence (e.g., signage) and licensed in-home daycares. Each unique entity (organization or business) at each unique address was considered an asset. All services or programs originating from a single entity at one address constituted a single asset. Three examples are provided here to further illustrate how we defined an asset: (1) multiple restaurants of a single franchise located at unique addresses were listed as separate assets; (2) a business office for a franchise was considered a unique asset if located at a different address from the restaurant or at the same address but with a different suite number; (3) a community center with many programs offered at a single address was considered a single asset.

To test the performance of the data collected, an initial list of assets was created from two main sources using the most current available data: a purchased Dun and Bradstreet¹⁴ list downloaded May 2009 and a list of nonprofit organizations from the National Center for Charitable Statistics (NCCS)³⁷ downloaded May 2009. Additional assets were added from a government list of licensed in-home daycares (State of Illinois, Department of Child and Family Services, May 2009), the government “blue pages” in the 2009 telephone directory (e.g., post offices), 2009 lists from the City of Chicago (e.g., public schools), and the university’s 2009 campus directory.

Prior to the start of field work, the initial list of assets was prepared, or “precleaned,” using the following steps: (1) each asset was assigned a unique six-digit ID, (2) assets with zip codes outside Chicago’s South Side were deleted, (3) P.O. boxes were deleted, (4) all addresses were standardized according to the United States Postal Service system, (5) a census tract location variable was added to each asset on the list, and (6) based on census tract location, all assets outside the six community pilot region were deleted. Lastly, duplicate listings, which included assets with a name and/or address that was the same or differed only typographically (e.g., two listings at the same address, “ymca” and “YMCA”) were deleted from the list.

To be conservative, multiple assets listed at a single address with different asset names were retained on the initial list, to be verified by direct observation in the field as detailed below. Because the field work was used to verify each asset on the initial list, preparation was different from what would have been done if the initial list were being prepared for secondary data analysis (for example, see Rundel et al.¹³).

Data Collection Tools

A custom taxonomy was built to index each asset by its primary function. It was designed to address key research questions (e.g., How many grocers are in the community?), populate an online database that generates customizable asset maps, and allow for easy, systematic data entry. Existing taxonomies did not meet these objectives; they were either sector-specific (e.g., National Taxonomy of Exempt Entities,³⁸ North American Industry Classification System³⁹), too detailed (e.g., Alliance of Information and Referral System taxonomy⁴⁰), or not detailed enough (e.g., Standardized Industrial Classification⁴¹). The custom two-level system (15 mutually exclusive core categories with 4 to 24 mutually exclusive subcategories) was informed primarily by the widely used Standardized Industrial Classification⁴¹ and National Taxonomy of Exempt Entities³⁸ taxonomies. Categories were iterated through team discussion and field testing.

Minimum data recorded for each observed asset included: (1) street address, (2) name, (3) taxonomy classification, and (4) disposition. Disposition was classified as “operating” (observed in the field and operating), “defunct” (closed, not found, duplicate, or ineligible), or “unknown” (observed in the field but operating status unknown). Assets identified as operating or unknown were also classified as “residential” (in a private home) or “nonresidential” (in a commercial or nonprofit space). When possible, phone number, website, and the name of owner, leader, or manager were collected.

Data were collected using web-enabled cell phones (Sprint™, HTC Touch Pro). The custom software, *MapApp*™, built expressly for this project, consisted of a web front end and a Perl (<http://www.perl.org>) and PostgreSQL (<http://www.postgresql.org/>) back end. The development and functional testing of *MapApp*™ occurred during research meetings and interaction between the census workers and the developers. The initial list of assets was uploaded to the *MapApp*™ database (henceforth referred to as the “prepopulated list”).

MapApp™ includes a data entry interface that allows census workers to search for, add, or update assets by address, intersection, or name. A unique interface for supervisors displays data useful for tracking field work (e.g., real-time summaries of data collected by each census worker) and screens to facilitate data entry. Data entry screens show a complete history of the data for each asset, crucial to tracking modifications to the prepopulated list, for example, correction of a listed asset name based on field observation. A dataset export and import system enables supervisors to edit multiple assets simultaneously. *MapApp*™ was designed for application across geographic regions and can accommodate local address conventions.

Census Worker Recruitment and Training

Census workers included paid teams of university students and community residents. University students were recruited primarily from Survey Lab employees because of their familiarity with the aims of basic research. Community residents were recruited primarily through the City of Chicago’s Youth Ready Program⁴² because of their key local knowledge. Of the 23 census workers employed, 11 were university students and 13 were community residents (groups not mutually exclusive).

Census workers received 15 h of training, including classroom and field sessions, using the cell phones and *MapApp*™. Training emphasized the use of the taxonomy and disposition variables and safety. Census workers wore matching project-logo shirts to enhance team spirit, foster community awareness, and promote safety.

Data Collection and Validation

Three-person census teams were deployed on foot for 4-h shifts; one census worker entered data, one provided walking directions and referenced field materials, and one fielded questions from bystanders and distributed brochures. As the workers grew confident with the data collection methods, two census workers per team entered data. In residential areas, teams traveled by car; one census worker drove, one entered data, and one provided driving directions and referenced field materials.

Data collected in the field were obtained from publicly posted signage and asset representatives (e.g., business owners or managers). Every block in the six communities was covered. During field work, census workers attempted to verify every asset on the prepopulated list. Census workers assigned dispositions to each asset on the prepopulated list and modified information as necessary (e.g., corrected the address, corrected typographical errors). Census workers also added newly identified assets observed in the field to the list.

A match between the prepopulated list and direct observation did not require a perfect asset name match or a perfect address match. Census workers were trained to understand that assets on the prepopulated list might be listed under a legal name that differs from a name used on signage and they should use all possible sources, including asset representatives, phone numbers, websites, representatives of neighboring assets, and community residents, to determine if the asset identified by direct observation “matched” that on the prepopulated list. Census workers were also trained to be aware of assets that might have more than one address, including assets located at the intersection of two streets and assets with multiple entrances. In all instances in which a “match” between the prepopulated list and direct observation was identified, the disposition was coded as “operating” if the asset was operating and “defunct” if the asset was listed and was no longer operating. Census workers were also trained to categorize assets as “unknown” unless they were able to determine the operating status with high confidence.

After each shift, teams marked the completed blocks on laminated maps to monitor progress. A supervisor was available by phone during all shifts and monitored progress in real-time using *MapApp*TM.

Web searches and telephone calls were used to obtain information not found in the field, primarily phone numbers, email addresses, and websites which few assets posted publicly. Census workers and supervisors revisited some blocks to validate data accuracy and index assets not resolved during initial field visits or calls. Census workers also rechecked dense commercial areas and buildings with multiple assets (e.g., one building contained >300 assets). Some areas were revisited at different times of day or different days of the week to collect data during posted operating hours.

Data Dissemination

Data were publicly disseminated approximately 16 weeks after initiating data collection. Operating assets were geo-coded and data were uploaded to <http://www.SouthSideHealth.org> (in Spanish, <http://www.DondeEsta.org>). These custom-built mapping site leverages open-source software (including Google Maps Application Programming Interfaces)⁴³ to enable untrained users to find built assets in specified communities. Only assets with a public presence or permission provided by the owner or leader were posted online. The website was promoted at a community forum attended by 104 individuals (45 % community), regular research meetings, and digital

communications (e.g., newsletters, blog [<http://southsidehealth.wordpress.com/>], twitter [<http://twitter.com/sshvs>]).

Statistical Analysis

Built assets were described by disposition and source (prepopulated list or newly identified during field work). The sensitivity of the Dun and Bradstreet¹⁴ list and the prepopulated list overall for detecting operating, nonresidential built assets was calculated by disposition, community, and sector. A positive match between the prepopulated list and direct observation was defined as any asset at a given address that was identified as “operating” through direct observations regardless of modifications to the asset’s name (e.g., an asset listed on the prepopulated list under its legal name and operating at the same address using a different name on its signage would be a positive match) or modifications to the asset’s address. Positive predictive value was not calculated because the number of “false positives” on the prepopulated lists could not be accurately quantified for several reasons. The “defunct” category included nonoperating assets, duplicates, and ineligible assets, and the methodology did not generate sufficient information to further classify these assets. Likewise, the protocol did not generate data to determine the disposition for the large number of residentially located assets.

RESULTS

Operations and Timeline

Research team meetings began in October 2008 (see Figure 1). Through February 2009, the research team developed the asset census methodology and taxonomy with input from community members, some of whom later joined the advisory board. Community advisors first met in June 2009 and monthly thereafter; frequent project updates were emailed between meetings. The community education campaign was initiated in June and followed by data collection from July 2009 to August 2009. During the data collection period, the research team met weekly to review progress,

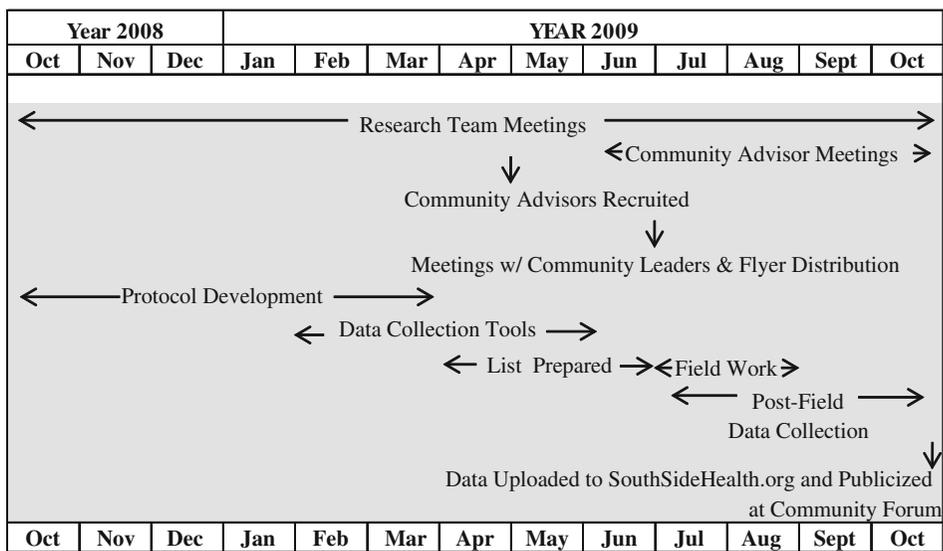


FIGURE 1. Timeline and key milestones for asset census project; Chicago, IL; 2008–2009.

resolve technical issues, and incorporate community advisors’ feedback. Data collection continued until the data were disseminated via <http://www.SouthSideHealth.org> in October 2009. In addition to operations staff and research team leaders, an average of 11 university team members regularly attended research meetings and an average of 7 community advisors attended advisory board meetings.

Community Education Campaign

Census workers spent approximately 432 person-hours distributing flyers in the 6 communities in advance of field work. During field work, census workers found that most community members they encountered were aware of the project, qualitatively suggesting that the campaign reached a large proportion of community members. No hostile encounters were documented; however, some community members asked whether the census was to inform real estate speculation by the university.

Data Collection Tools

MapApp™ functioned well. The cell phones were durable and had good connectivity. The customized taxonomy remained largely intact.

Community Assets

After preparation of the secondary data sources, the prepopulated list included 5,773 assets: 90 % from the Dun and Bradstreet¹⁴ list, 7 % from the NCCS³⁷ list, and the remainder from government lists. The asset census field work identified 801 additional assets. All identified assets were classified in the field as operating, defunct, or unknown (Figure 2). Of the total 6,574 assets (5,773 from the prepopulated list and 801 from field work), 2,265 were operating. Of all operating assets, 29 % (n=653) would have been missed if one relied solely on the prepopulated list generated from the secondary data sources using the preparation steps described in the “Methods” section.

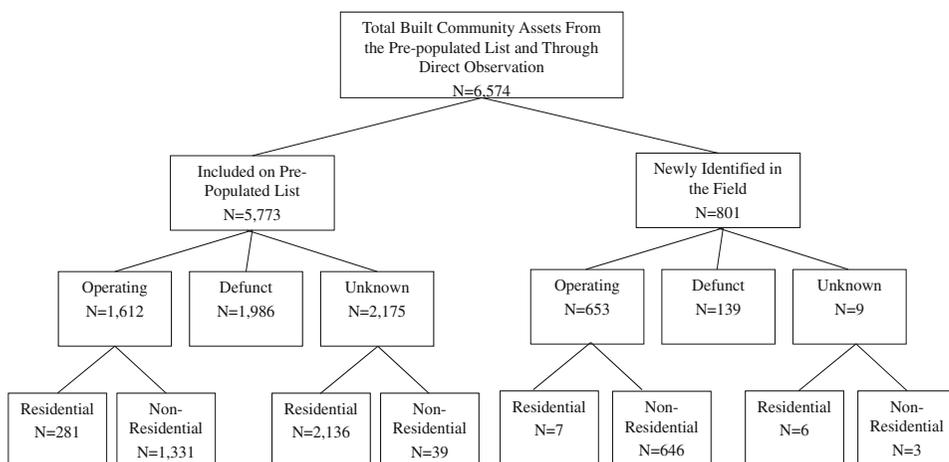


FIGURE 2. Final disposition of all assets; Chicago, IL; 2009. Each asset was indexed as one of the following dispositions: operating (observed in the field and operating), unknown (observed in the field but operating status unknown), or defunct (not operating, not observed in the field, or not eligible base on definition of an asset). Operating and unknown assets were also categorized as nonresidential (physically located in a building primarily used as a commercial space) or residential (physically located in a building used primarily as a private residence). It was not possible to determine if defunct assets were residential or nonresidential.

More than 34 % ($n=1,986$) of the assets on the prepopulated list were categorized as defunct through direct observation. This included assets that were closed, not found, duplicate, or ineligible based on field observation; these assets could not be further classified using our methodology. The remaining 2,175 assets on the prepopulated list were classified as unknown—census workers could not determine if these assets were operating or defunct. The vast majority of unknown assets ($n=2,136$, 98 %) had no street presence and were located in residential spaces. In an effort to resolve assets classified as unknown during field work, telephone calls and web searches were used. The list included phone numbers for 52 % of the assets classified as unknown; 46 % of these were nonworking numbers.

The majority of operating assets ($n=1,977$, 83 %) were located in nonresidential spaces. For operating assets located in nonresidential space, the Dun and Bradstreet list¹⁴ had a sensitivity rate of only 61 % (Table 1) using direct observation as the criterion standard. Combining Dun and Bradstreet¹⁴ with other secondary data sources increased sensitivity for operating assets to only 67 %. Sensitivity of the prepopulated list varied qualitatively by community (range, 64–73 %) as well as by sector type (ranging from 60 % for wholesale, storage, and transit to 81 % for religious assets). Sensitivity of the prepopulated list for the health services sector was relatively low (64 %), but slightly better for the human services sector (72 %); for the Dun and Bradstreet list alone, the sensitivity for health and human services was 62 and 57 %, respectively. The sensitivity of the prepopulated list for each sector also varied by community (data not shown). For example, for the dining sector in the community of Hyde Park, the Dun and Bradstreet¹⁴ list had a sensitivity of 66 % compared to only 55 % in neighboring Grand Boulevard. For the retail sector, the Dun and Bradstreet¹⁴ list had a sensitivity of 58 % in Hyde Park versus 80 % in Grand Boulevard.

For operating assets, 48 and 34 % of asset names provided by Dun and Bradstreet¹⁴ and NCCS, respectively, required edits (e.g., capitalization, expanding abbreviations). More than 80 % of the asset names from government lists required edits (e.g., in-home daycares were listed by provider name, but many providers used other names for their business, the city used abbreviations like BRCH for Branch).

Based on targeted field, web, and telephone rechecks after the initial field work, data were added or modified for 2,835 of the 6,574 assets. For the vast majority of assets, nonrequired data (e.g., telephone numbers) were added; only 6 % required address edits, 8 % required disposition edits, and 4 % required taxonomy classification edits.

It took 52 days and 900 person-hours to cover all 2,602 face blocks in the 6 communities (8–22 days per community); 11.8 % of blocks were mostly commercial. An additional 1,700 person-hours were spent on post-field work (320 census worker/supervisor hours and 1,380 research assistant hours). In total, approximately 3,000 person-hours were spent on flyering, data collection, and post-field work.

Data Dissemination

Since the public site with the asset census data was launched October 2009, the number of <http://www.SouthSideHealth.org> users has steadily grown (approximately 40,000 visitors and 986,000 hits as of January 2012). A broad spectrum of community and university entities have requested customized data analyses and access to raw asset census data (for examples, see Table 2).

TABLE 1 Source of assets by final disposition, community area, and sector; Chicago, IL; 2009

	Total assets		Asset on prepopulated list				Newly identified assets		Sensitivity of prepopulated list for nonresidential operating assets	Sensitivity of Dun & Bradstreet list for nonresidential operating assets	
	n	n	Dun and Bradstreet		NCCS		Government ^a				
			(%)	(%)	n	(%)	n	(%)			
Final disposition ^b											
Operational, nonresidential	1,977	1,206	(61.0)	64	(3.2)	61	(3.1)	646	(32.7)	67%	61%
Operational, residential	288	196	(68.1)	3	(1.0)	82	(28.5)	7	(2.4)		
Unknown, nonresidential	42	29	(69.0)	10	(23.8)	0	(0.0)	3	(7.1)		
Unknown, residential	2,142	1,958	(91.4)	155	(7.2)	23	(1.1)	6	(0.3)		
Defunct	2,125	1,796	(84.5)	142	(6.7)	48	(2.3)	139	(6.5)		
Community area ^c											
East Side	302	198	(65.5)	4	(1.3)	16	(5.3)	84	(27.8)	70.6	67.5
Grand Boulevard	442	276	(62.4)	12	(2.7)	40	(9.0)	114	(25.8)	70.6	63.7
Hyde Park	779	484	(62.1)	31	(4.0)	10	(1.3)	254	(32.6)	64.0	58.9
Kenwood	174	111	(63.8)	3	(1.7)	21	(12.1)	39	(22.4)	66.7	52.6
Washington Park	185	115	(62.2)	7	(3.8)	18	(9.7)	45	(24.3)	73.1	64.1
Woodlawn	383	218	(56.9)	10	(2.6)	38	(9.9)	117	(30.5)	64.9	57.9
Sector ^d											
Arts and entertainment	78	52	(66.7)	2	(2.6)	0	(0.0)	24	(30.8)	64.7	61.8
Childcare and schools	336	103	(30.7)	9	(2.7)	134	(39.9)	90	(26.8)	63.3	38.4
Dining	172	105	(61.0)	0	(0.0)	0	(0.0)	67	(39.0)	61.1	61.1
Financial, insurance, and real estate	330	244	(73.9)	1	(0.3)	0	(0.0)	85	(25.8)	64.8	64.5
Fitness	26	18	(69.2)	1	(3.8)	0	(0.0)	7	(26.9)	70.8	66.7
Health services	147	93	(63.3)	4	(2.7)	0	(0.0)	50	(34.0)	63.8	61.6
Industrial	12	8	(66.7)	0	(0.0)	0	(0.0)	4	(33.3)	60.0	60.0
Personal service	179	120	(67.0)	0	(0.0)	0	(0.0)	59	(33.0)	66.5	66.5
Programmed residential	77	52	(67.5)	2	(2.6)	0	(0.0)	23	(29.9)	69.4	66.7

Public service	46	21	(45.7)	1	(2.2)	7	(15.2)	17	(37.0)	63.0	45.7
Religious	213	145	(68.1)	27	(12.7)	0	(0.0)	41	(19.2)	80.7	68.4
Retail	287	192	(66.9)	0	(0.0)	0	(0.0)	95	(33.1)	66.7	66.7
Human services	133	80	(60.2)	17	(12.8)	1	(0.8)	35	(26.3)	72.3	57.1
Trade service	164	128	(78.0)	2	(1.2)	0	(0.0)	34	(20.7)	73.0	72.2
Wholesale, storage, and transit	46	29	(63.0)	0	(0.0)	1	(2.2)	16	(34.8)	60.0	57.5

^aIncludes lists from the State of Illinois and City of Chicago

^bEach asset was indexed as one of the following dispositions: operating (observed in the field and operating), unknown (observed in the field but operating status unknown), or defunct (not operating, not observed in the field, or not eligible base on definition of an asset). Operating and unknown assets were also categorized as nonresidential (physically located in a building primarily used as a commercial space) or residential (physically located in a building used primarily as a private residence). Population totals obtained from 2010 US Census. Percent of households living at or below 200 % of the federal poverty level calculated using 2005–2009 American Community Survey data

^cExcludes assets with dispositions of unknown or defunct

^dExcludes assets with dispositions of unknown or defunct, and 19 other assets (5 for which taxonomy could not be determined and 14 which did not fit in the 15 categories and were indexed as other)

TABLE 2 Example asset census data use cases; Chicago, IL; 2009

Organization	Use
Community advocacy group	Using maps of community children's health clinics to support health component of a federal grant application
Community organization	To create an internal tracking system to better recruit and vet developments, potential businesses, and service providers against existing assets and needs
University and Department of Public Health	To examine existing, community-level resources related to prevalence of preventable diseases and deaths
Community leader	Illustrate asset needs within the community in order to attract new businesses
Community leader	To support a proposal to collect information on faith-based organizations
Corporation	To understand availability of health-related resources such as grocers and fitness centers
University student program	To understand the assets available in the community in order to inform their community service projects
University social scientists	To find community-based organizations to interview for research study exploring the extent to which these organizations believe they represent the community
University physician	To find primary care physicians to interview for research study
State-funded program	To enrich database for case managers and social workers, especially services available for children

DISCUSSION

This project is part of a larger effort aimed at improving the health of low-income urban populations and optimizing the built urban context to improve the likelihood that “individuals’ default decisions are healthy.”⁴⁴ Using an asset-based community-engaged research approach, this team is developing methodologies to measure the built environment that can be replicated in other communities.²⁹ These data are being used to better understand relationships between the built asset environment and health and translated to improve health care delivery and community-based referral (see “CommunityRx system: linking patients and community-based service” at <http://innovations.cms.gov/initiatives/Innovation-Awards/illinois.html>).

This community–university research team developed methodology to conduct a census of all public-facing, built assets via direct observation. Using data generated from this methodology, we tested the sensitivity of secondary data sources commonly used for health research. Census workers identified 2,265 operating built assets via street-level observation. More than a third of the assets listed in the prepared secondary datasets were categorized during field work as defunct (closed, not found, duplicate, and ineligible). For operating assets located in nonresidential spaces, the most up-to-date Dun and Bradstreet¹⁴ data available had a sensitivity rate of 61 %; in other words, 39 % of the built assets that were identified via direct observation in the field were not accounted for in the secondary datasets.

This sensitivity rate was slightly higher than that found by Lake et al. for built food assets using an online directory and the yellow pages (51 and 53 %, respectively), but lower than that Lake et al. found using UK government data (84 %).²⁴ Sensitivity was lowest for sectors expected to have higher turnover (e.g., dining) compared to those expected to have less turnover (e.g., religious institutions). Variations in the sensitivity

by sector, community, and sector among communities suggest possible systematic differences in the sensitivity of secondary datasets by community and sector, indicating that secondary datasets may not be a reliable source of asset data for the purpose of these comparisons, even with extensive cleaning.

Previous work assessing the associations between variation in the built asset environment and health has relied on secondary data sources^{7,9-13,15-21} and the potential impact of error in the measurement of built assets on these studies' results is not well understood. Sensitivity for health and human services assets was abysmal, with more than one third of health services and one quarter of human services missing from the secondary datasets. Where secondary datasets are used to inform referrals, missing information impairs access to community assets and could perpetuate a deficit-based view of the community. Inclusion of defunct assets also presents a costly problem for health care providers offering community-based referrals and for the patients receiving them. The consequences of referring patients to nonoperational assets include frustration, despair, erosion of trust, delayed treatment, and loss of scarce resources like time and money required to seek help. Accurate, real-time knowledge about the built asset environment is fundamental to improving urban population health.

The number of unique users and site hits on <http://www.SouthSideHealth.org> and requests for raw data indicate broad value of the built asset data to researchers, community leaders, and other stakeholders. Data are being used to find assets needed to prevent and manage disease, map the strengths and weaknesses of the service delivery system, improve effectiveness of referrals, support growth and advocacy efforts of community development organizations, and provide input critical for urban planners and policymakers.

There are several limitations to this methodology. Census workers may have missed assets or incorrectly indexed assets; however, our validation efforts (e.g., revisiting dense commercial areas) mitigate this bias. Identification of false positives was limited because the "defunct" code aggregated closed, not found, duplicate, and ineligible assets. The steps to prepare the secondary datasets were conservative with the intent that field work would serve to verify any discrepancies that could not be readily resolved in the lab. Different data preparation steps would be needed to prepare a secondary data source for, as an example, ranking neighborhoods by food asset availability. The secondary datasets we used, applying different preparation methods, could be compared to the dataset we generated in the field for the purpose of rank or other comparisons; these analyses were beyond the scope of the current work.

Additionally, this census represents a single point in time and occurred during a time when much of the country was experiencing an economic crisis. This could bias the results if the economic crisis resulted in an above-average turnover in assets, which would in turn suggest that the sensitivity of the prepopulated list was lower than is typical. Longitudinal comparisons between secondary datasets and data generated using this asset census methodology would be necessary to better understand this potential bias. Data collection did not include granular data on all goods and services available within each built asset. For example, a church was identified exclusively as such; other services provided by the church (e.g., food pantry) were not identified. The ability to index residentially located assets was limited because the methodology was designed to capture public-facing assets; few residentially located assets on the prepopulated list had signage. The large number of residentially located assets on the prepopulated list suggest that these assets, as well

as those outside the formal economy (“off the books”),⁴⁵ may represent a large segment of the local economy. Finally, we are not able to compare costs across methods due to the proprietary nature of secondary data sources. However, this method provides paid jobs and STEM-type training⁴⁶ in communities where jobs are scarce and is building a bridge between university researchers and community leaders. The societal value of this strategy should be considered in evaluating relative return on investment.

Our novel methodology advances public health efforts to understand the impact of built assets on health and to create systems to better link people to health-promoting community assets. Pairing these directly observed, empirical data on the built asset environment with real-time, granular health data will shed light on the mechanisms by which the built asset environment influences health. The community-engaged strategy facilitated meaningful ongoing community–university collaborations that are critical to sustaining and growing the effort. In addition, the community-engaged approach laid a strong foundation for implementing longitudinal work that includes expansion of the geographic region, continuous updating of the data, and a prospective study of population health and the built asset environment.

ACKNOWLEDGMENTS

We would like to acknowledge the invaluable input from our community partners who guided and co-led us in the development and implementation of the asset census methodology. We would also like to acknowledge the efforts of literacy expert Shane Desautels, software engineer Evgeny Selkov, graphic designer Jola Glotzer, and GIS specialist Todd Schuble as well as Scott Allard and Jennifer Mosley for their significant input on the use of secondary datasets and data collection methodologies. We would also like to acknowledge the efforts of the South Side Health and Vitality Studies staff who provided ongoing support to this project and without which none of this would have occurred. This project was supported by the South Side Health and Vitality Studies (P.I. Stacy Lindau, MD, MAPP). The South Side Health and Vitality Studies are supported by funding from the University of Chicago Medical Center Division of the Biological Sciences; the Office of the Urban Health Initiative; the Walter G. Zoller Memorial Fund at the University of Chicago; the Chicago Community Trust; the George Kaiser Family Foundation; the Otho S. A. Sprague Memorial Institute; Patricia O. Cox, and the National Institute on Aging at the National Institutes of Health (K23AG032870 and 1RC4AG039176-01). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging or the National Institutes of Health.

REFERENCES

1. Franks P, Clancy CM, Gold MR. Health insurance and mortality. *JAMA*. 1993; 270: 737–741.
2. Institute of Medicine. *For the Public's Health: The Role of Measurement in Action and Accountability*. Washington: The National Academies Press; 2010.
3. Lurie N, Dubowitz T. Health disparities and access to health. *JAMA*. 2007; 297(10): 1118–1121.
4. Marmot M, Wilkinson RG, eds. *Social Determinants of Health*. Oxford: Oxford University Press; 1999.

5. McGinnis JM, Foege WH. Actual causes of death in the United States. *JAMA*. 1993; 270(18): 2207–2212.
6. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. *JAMA*. 2004; 291(10): 1238–1245.
7. Block JP, Christakis NA, O'Malley AJ, Subramanian SV. Proximity to food establishments and body mass index in the Framingham Heart Study Offspring Cohort over 30 years. *Am J Epidemiol*. 2011; 174(10): 1108–1114.
8. Boone-Heinonen J, Gordon-Larsen P, Kiefe CI, Shikany JM, Lewis CE, Popkin BM. Fast food restaurants and food stores: longitudinal associations with diet in young to middle-aged adults: the CARDIA study. *Arch Intern Med*. 2011; 171(13): 1162–1170.
9. Dubowitz T, Ghosh-Dastidar M, Eibner C, et al. The Women's Health Initiative: the food environment, neighborhood socioeconomic status, BMI, and blood pressure. *Obesity*. 2011; 20: 862–871. doi:10.1038/oby.2011.141.
10. Gordon-Larsen P, Nelson MC, Page P, Popkin BM. Inequality in the built environment underlies key health disparities in physical activity and obesity. *Pediatrics*. 2006; 117(2): 417–424.
11. Li F, Harmer P, Cardinal BJ, et al. Built environment and 1-year changes in weight and waist circumference in middle-aged and older adults. *Am J Epidemiol*. 2009; 169: 401–408.
12. Moreland K, Wing S, Diez Roux A. The contextual effect of the local food environment on residents' diets: the atherosclerosis risk in communities study. *Am J Public Health*. 2002; 92(11): 1761–1767.
13. Rundle A, Neckerman KM, Freeman L, et al. Neighborhood food environment and walkability predict obesity in New York City. *Environ Health Perspect*. 2008; 117(3): 442–447.
14. Dun and Bradstreet. Dun and Bradstreet. 2009; <http://www.dnb.com/>. Accessed January 4, 2011.
15. Grafova IB, Freedman VA, Kumar R, Rogowski J. Neighborhoods and obesity in later life. *Amer J Public Health*. 2008; 98(11): 2065–2071.
16. Burdette HL, Whitaker RC. Neighborhood playgrounds, fast food restaurants, and crime: relationships to overweight in low-income preschool children. *Prev Med*. 2004; 38(1): 57–63.
17. Gibson DM. The neighborhood food environment and adult weight status: estimates from longitudinal data. *Am J Public Health*. 2011; 101(1): 71–78.
18. Kelly B, Flood VM, Yeatman H. Measuring local food environments: an overview of available methods and measures. *Health Place*. 2011; 17(6): 1284–1293.
19. Maddock J. The relationship between obesity and the prevalence of fast food restaurants: state-level analysis. *Am J Health Promot*. 2004; 13(2): 137–143.
20. Nelson MC, Gordon-Larsen P, Song Y, Popkin BM. Built and social environments associations with adolescent overweight and activity. *Am J Prev Med*. 2006; 31(2): 109–117.
21. Wang MC, Cubbin C, Ahn D, Winkleby MA. Changes in neighbourhood food store environment, food behaviour and body mass index, 1981–1990. *Public Health Nutr*. 2008; 11(9): 963–970.
22. Bader MD, Ailshire JA, Morenoff JD, House JS. Measurement of the local food environment: a comparison of existing data sources. *Am J Epidemiol*. 2010; 171(5): 609–617.
23. Boone JE, Gordon-Larsen P, Stewart JD, Popkin BM. Validation of a GIS facilities database: quantification and implications of error. *Ann Epidemiol*. 2008; 18(5): 371–377.
24. Lake AA, Burgoine T, Greenhalgh F, Stamp E, Tyrrell R. The foodscape: classification and field validation of secondary data sources. *Health Place*. 2010; 16: 666–673.
25. InfoUSA. InfoUSA. 2011; <http://www.infousa.com>. Accessed November 8, 2011.
26. Commission on Social Determinants of Health. *Closing the Gap in a Generation: Health Equity through Action on the Social Determinants of Health*. Geneva: World Health Organization; 2008.
27. Robert Wood Johnson Foundation. 2011. Health Care's Blind Side. <http://www.rwjf.org>. Accessed July 20, 2012.

28. Hill LD, Madara JL. Role of the urban academic medical center in US health care. *JAMA*. 2005; 294(17): 2219–2220.
29. Lindau S, Makelarski J, Chin M, et al. Building community-engaged health research and discovery infrastructure on the South Side of Chicago: science in service to community priorities. *Prev Med*. 2011; 52(3–4): 200–207.
30. Kretzmann JP, McKnight JL. *Building Communities from the Inside Out: A Path toward Finding and Mobilizing a Community's Assets*. Evanston: ACTA Publications; 1997.
31. Santill A, Carroll-Scott A, Wong F, Ickovics J. Urban youths go 3000 miles: engaging and supporting young residents to conduct neighborhood asset mapping. *Am J Public Health*. 2011; 101(12): 2207–2210.
32. Lindau S, James R, Makelarski J, Sanders E, Johnson D. Comments from the South Side of Chicago on New Haven's Inspiring Initiative. *Am J Public Health*. 2012; 102(7): e3–e4.
33. City of Chicago. Explore Chicago. 2011. <http://www.explorechicago.org/city/en/neighborhoods.html>. Accessed August 30, 2011.
34. U.S. Census Bureau. 2010 U.S. Census Data. 2010. <http://www.census.gov>. Accessed 3 January 2011.
35. U.S. Census Bureau. 2005–2009 American Community Survey 5-Year Estimates. 2010. <http://www.americanfactfinder.gov>. Accessed 3 January 2011.
36. Lee W. Chicago South Side Mapping Project Shows Neighbors Live Worlds Apart: Students Documenting Health-Care Shortcomings—Including Lack of Healthy Foods, Medical Clinics—in Woodlawn Neighborhood. *The Chicago Tribune*. 2009.
37. Urban Institute. National Center for Charitable Statistics. 2009; <http://nccs.urban.org/>. Accessed May 1, 2009.
38. National Center for Health Statistics. National Taxonomy of Exempt Entities. 2009; <http://nccs.urban.org/classification/NTEE.cfm>. Accessed January 3, 2010.
39. U.S. Census Bureau. North American Industry Classification System (NAICS). <http://www.census.gov/eos/www/naics/>. Accessed January 4, 2011.
40. Information and Referral Federation of Los Angeles County. The AIRS/211 LA County Taxonomy of Human Services. 2008; <http://www.211taxonomy.org>. Accessed February 7, 2011.
41. U.S. Department of Labor. Standard Industrial Classification (SIC) System. http://www.osha.gov/pls/imis/sic_manual.html. Accessed January 4, 2011.
42. Youth Ready Chicago. 2010; <http://www.youthreadychicago.org>. Accessed February 2, 2011.
43. Google Maps Application Programming Interfaces. 2009. <https://developers.google.com/maps/> Accessed October 2009.
44. Frieden TR. Framework for public health action: the health impact pyramid. *Am J Public Health*. 2010; 100(4): 590–594.
45. Venkatesh S. *Off the Books: The Underground Economy of the Urban Poor*. Cambridge: Harvard University Press; 2006.
46. STEM Education Coalition. 2011. <http://www.stemedcoalition.org>. Accessed November 4, 2011.