Creating internationally-comparable built environment variables in GIS for the IPEN studies

Marc A. Adams
Arizona State University/
IPEN Coordinating Center

GHENT 2013

Study sequence flowchart

1. Prepare → **2. Collect** → **3. Manage** →

- Obtain GIS data
- Neighborhood selection
- Survey translations/approvals
- Accelerometers & training
- Years 1-2 funding

- Recruitment & monitoring
- Survey checking
- Accelerometer compliance screening
- Years 1-2 funding

- File & variable names
- Survey cleaning
- GIS templates
- CC comparability assessments
- Years 2-3 funding

4. Process → **5. Papers** → **6. Policy!**

- GIS buffers (countries)
- Surveys (CC)
- Accelerometers (CC)
- Years 2-3 funding

- Select topics
- Proposals for review with Publications Committee
- Analysis
- Manuscripts
- Year 4 funding

- Communicate results to local, national & international policy makers
- Create summary briefs
- Start making contacts
- Year 4+ funding

Overview

1. Discuss the process of developing a robust dataset of GIS-derived measures.

2. Review IPEN GIS Templates

3. Introduce variable naming conventions

4. Review quality control and comparability evaluations

5. Discuss future tools and needs

High Level View of GIS process

- Identify participants and their residences

- Geocode and create individual-level buffers

- Geoprocess spatial data for walkability components, recreation environment, transit, etc. variables

- Export variables to spreadsheet and name variables using naming conventions

- Ensure quality of data before sending to IPEN CC

Why GIS Coordination is needed?

- To develop a common protocol for GIS variables including a set of operational definitions for walkability components, recreation, transport environments

- To provide *required* and *desired* **variables** and *recommended* and *acceptable* **procedures**

- To provide support and solutions to problems that generalize across countries

- To document GIS decisions and judge comparability of resulting variables across countries

IPEN Adolescent Countries

Australia

Brazil

Hong Kong

New Zealand

Switzerland

Bangladesh

Czech Republic

Israel

Nigeria

Taiwan

Belgium

Denmark

Japan

Portugal

United States

Germany

Malaysia

Spain

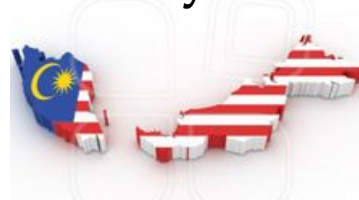**IPEN Ghent 2013**

GIS Coordination

- Two teams to help you!

- IPEN Coordinating Center
  - Marc Adams (GIS coordinator / Templates)
  - Kelli Cain
  - Terry Conway
  - Jacqueline Kerr
  - Jim Sallis

- Urban Design for Health (UD4H)
  - Larry Frank
  - Jim Chapman
  - Provide expertise, consulting and technical trainings

GIS Coordination Cont.

- GIS memo meeting in April 2013
  - Made some key decisions

- IPEN Adolescent GIS Templates
  - Drafts sometime between July-Sept 2013
  - Drafts will be open to countries for comments, input and feedback
  - Final versions available afterwards

- Ongoing support afterwards

Individual Level Buffers

- Right buffer sizes and types are key to good start

- **Required**: 500m and 1 km Street Network

- **Desired**: to be determined by end of summer
  - Residential candidates include:
    - 250m, 2km street network,
    - 250m, 500m, 1km, 2km pedestrian-enhanced street network
  - Buffers around child's school
    - Match to residential sizes and types
  - Routes to child's school

- Check with Marc before starting

Individual Level Buffers Cont.

- GIS Memo call revealed more info needed on sausage buffer and pedestrian-enhanced approaches

- UD4H planning to evaluate differences between sausage, detailed and generalized buffers
  - Suzanne Mavoa (AUS/NZ) plans to compare methods with IPEN adult data
  - Perhaps work together
  - One more country?

- Pedestrian-enhanced vs. street network
  - Comparison with adult data possible

GIS Measures

- **Hierarchy of possible GIS measures**

- **Required** Variables
  - Variables judged to be most available and likely across countries
  - Procedures least intensive
    - Examples: land area measures, counts, distance, gross residential density
  - All countries should produce required variables, if information available

- **Desired** Variables
  - Variables judged to require higher resolution data
  - Procedures more intensive
    - Examples: building floor areas, net residential density, signalized intersections
  - Desired variables should be calculated **in addition** to the required variables
  - Desired to be determined

- Based on data availability: not a global decision
  - Decisions made for specific templates and variables

GIS Procedures

- **Recommended procedures** should be used **instead of** acceptable procedures
  - Recommended procedures judged to be more precise methods of calculating the variables
  - Example: apportion parcels if centroid falls within participant's buffer

- **Acceptable procedures** should be used if recommended procedures cannot be used, or if recommended procedures have been deemed inappropriate for country-specific reasons
  - Acceptable procedures are less accurate than desired procedures, but acceptable to use
  - Example: apportion parcels if any area of parcel falls within participant's buffer

Built Environment and Physical Activity:

GIS Templates and Variable Naming Conventions

For the IPEN Studies*

April 2012

Marc A. Adams, Ph.D.[1]

James Chapman, MSCE[2]

James F. Sallis, Ph.D.[1]

Larry D. Frank, Ph.D.[2]

& International Physical Activity and Environment Network (IPEN) Study Coordinating Center[1]

[1]University of California, San Diego

[2]Urban Design for Health, Inc.

*This comprehensive set of documents was created as part of the International Physical Activity and Environment Network (IPEN) Study. It attempts to provide a common set of built environment definitions and measurement procedures for investigators in the IPEN study. This set of documents represents an evolving product that will be updated as needed for the Adult and Adolescent studies.

Available at: www.ipenproject.org

Table of Contents

SECTION 10: INTERSECTION DENSITY

Aim:	Develop a standardized definition of **intersection density** to be compared across participants across countries.
Task:	To identify and count intersections on a walkable road network that are within participants' 500 and 1000 meter buffers. To divide the buffer-level intersection counts by the total buffer land area, thereby creating intersection density (counts per sq km).
Datasets:	Road network and participant buffers
Definition:	The same walkable road network used for buffer creation should be used here. Roads where pedestrians are prohibited such as freeways and ramps are removed from the network before intersections are identified and counted. For example, limited-access freeways, toll roads, on and off ramps to them, interchanges between these road types should be removed from the network. "Intersection" means a point where 3 or more walkable road segments intersect.
Sidenotes:	Some countries may need to buffer intersection points because some segments in the road network were not properly aligned. Investigators should decide whether spatial misalignment is a problem or not. If it is a problem, then each country should decide the appropriate buffer size after observing their data.

Details:	**Please respond to the questions below:**	Excluded /NO	Included /YES
64.	Was the same walkable road network used for buffer creation used for the creation of intersection density? If a different road network was used how was it different? What was the reason for using a different road network? Please describe below:		
65.	Were pseudo-nodes (i.e. nodes that split road segments at non-intersections) and cul-de-sacs removed from the road network		

Overall goal

Tasks to be completed

Datasets needed

Operational definitions

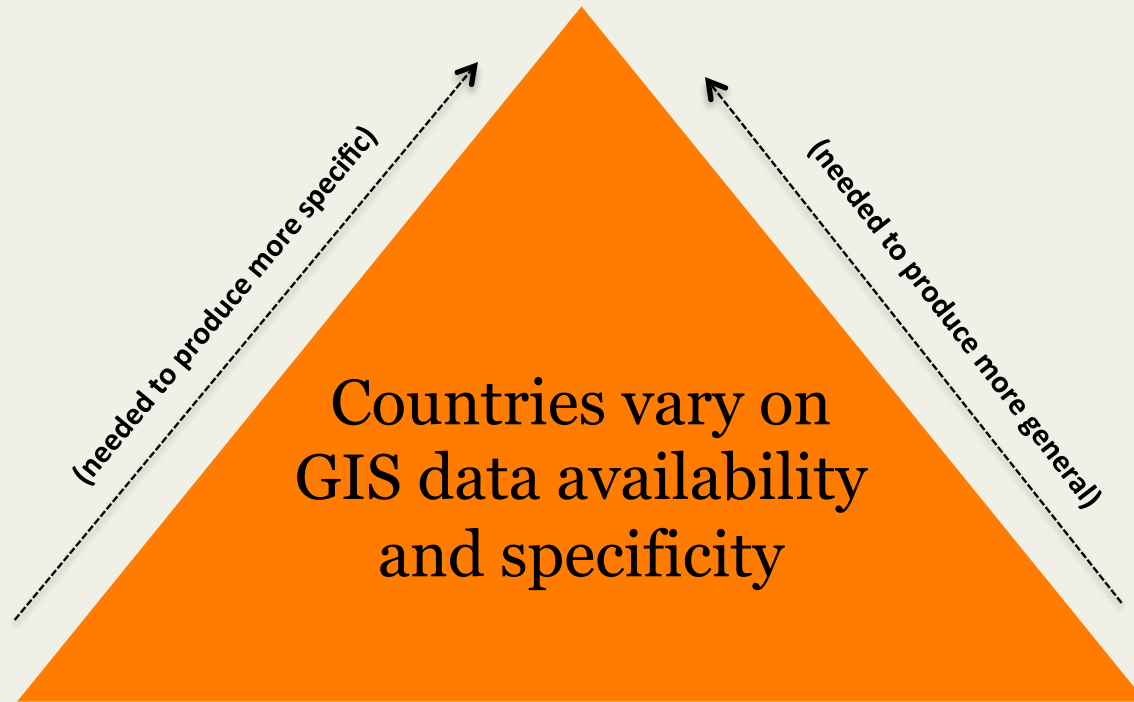Notes to be aware of...

Document work

IPEN Templates

- Let's review a few templates!

- Visit http://www.ipenproject.org/methods_gis.html

Comparable GIS Data

Least Common Denominator

(needed to produce more specific)

(needed to produce more general)

Countries vary on
GIS data availability
and specificity

Country's data
too crude or generalized

Country's data
too detailed or specific

Example: Land Uses

Figure 1: Land Cover Data (left) vs. Parcel Data (right). Source: Urban Design 4 Health., Inc.
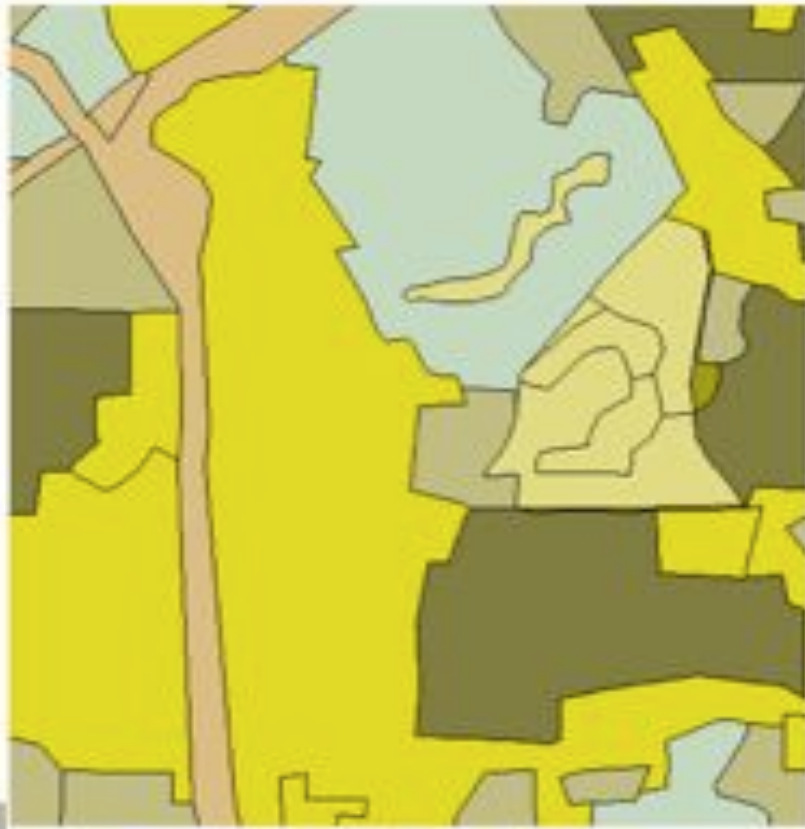
Generalized land use categories   Detailed land use categories

Example: Bus Stops

- Required: count of bus stops

- Bus routes available but <u>no formal stops available</u> in attributes

- Recommendation: create points at each intersection along bus route to make comparable

IPEN Adolescent: New Templates

- Neighborhood selection
- Buffer for route to child's school
- Buffers around child's school
- New Land Uses
  - School land use
  - Office land use
  - Vacant land use
- Transit level of service
- Regional accessibility

Variable Naming Conventions

- Lots of GIS variables possible
- Alphanumeric names
- Visual diagrams and key to help follow logic
- Bold = required variable
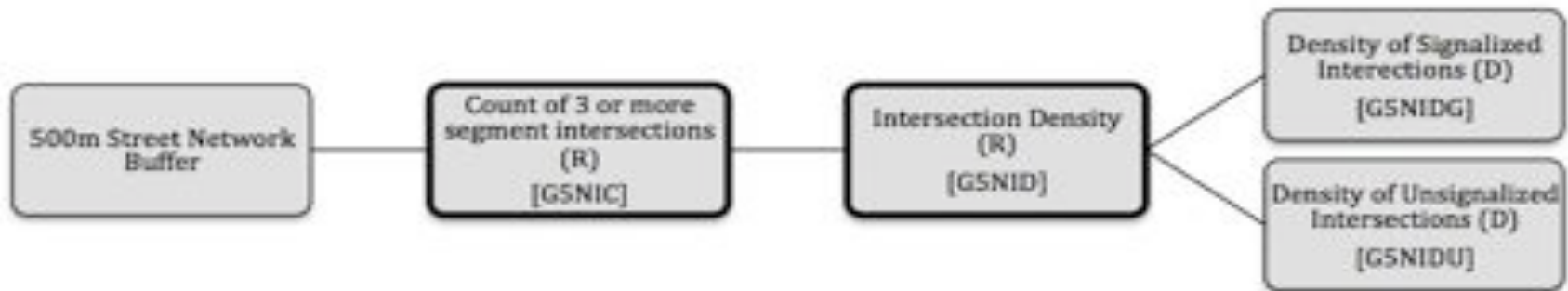
Variable Naming Conventions

- Organized by buffer size and type, and template
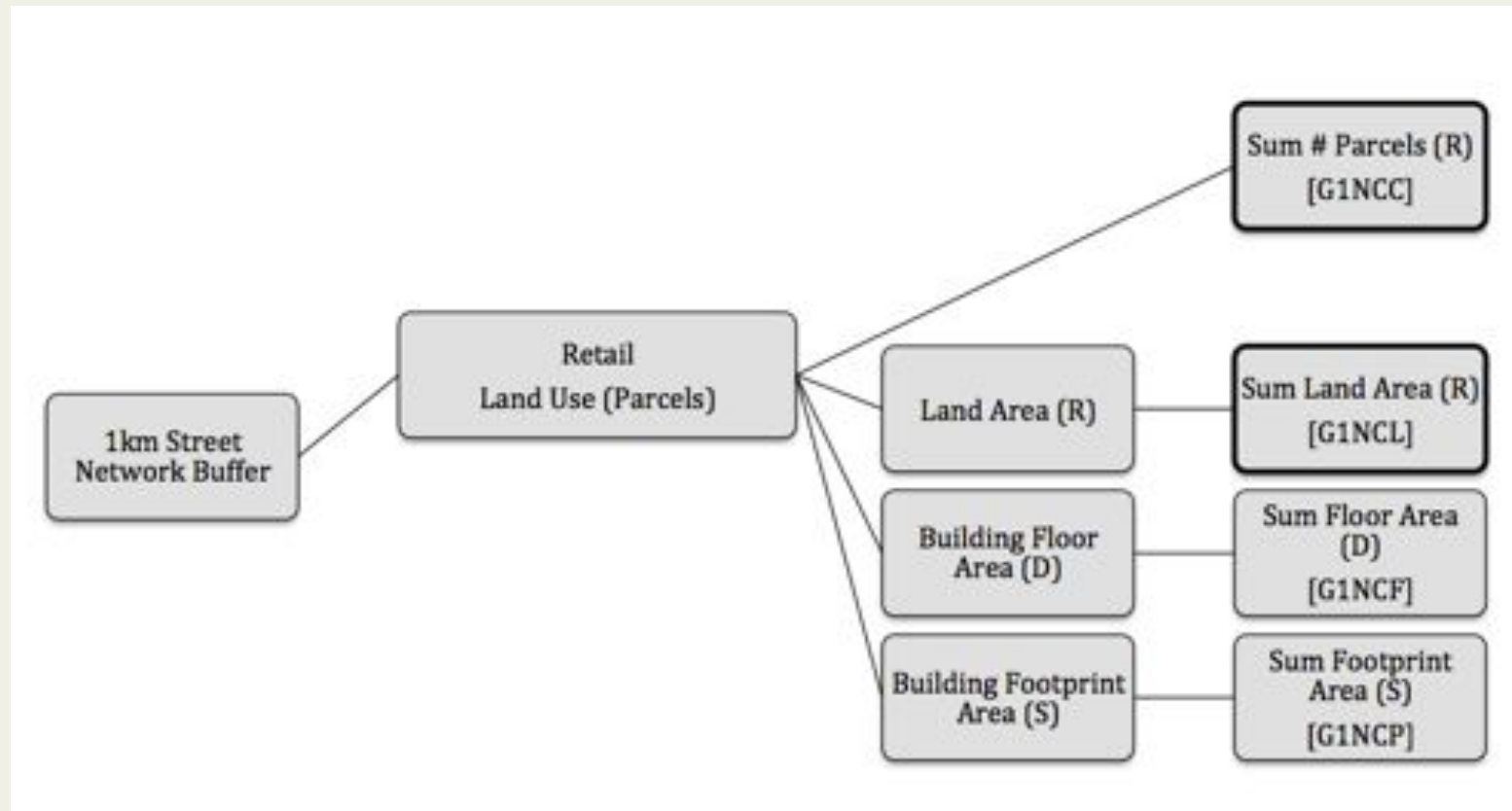
Variable Naming Conventions

- Participant IDs same as survey and accelerometer

- Make sure GIS analyst uses country code in participant IDs

- Name variables before sending based on convention documents and double check
  - CC limited ability to tell if done correctly!

Quality Control and Comparability Evaluation

- Templates are central to comparability evaluation

1. First send completed Templates to Marc at CC
   - Answer all Template questions
     - Each question must have answer (not applicable is okay)
   - Provide answers for all Templates, even for ones without data
     - Write "n/a" on first question of Template when no data available in your country

2. CC reviews templates for completeness (1-2 weeks)
   - Missing or incorrect information
   - Logic checks within and across templates
   - Preliminary review of comparability
   - Feedback on Templates and requests for more information

Quality Control and Comparability Evaluation

3. Once Templates approved, invited to upload data
   - Secured website or FTP

4. Preliminary data review (1-2 weeks)
   - Participant ID checks
   - Variable naming convention checks
   - Completeness (match Templates)
   - Logic and calculation checks on variables
   - Requests for missing data, miscalculated variables, etc.

4. Comparability evaluation by 2 raters (several months)

5. Further requests / new analyses based on comparability findings

6. Inclusion in pooled dataset

Arriving at Destination

- Internationally-comparable GIS variables

✓ Passed Comparability Evaluation

✓ Cleaned GIS datasets with proper IDs

✓ Data dictionaries complete and useful

	AUS	BEL	BRA	COL	CZ	DEN	HK	MEX	NZ	UK	US
Net Residential Density	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Intersection Density	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Retail Parcel Count	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Park Density	✓	✓	✓	✓	✓	✓	✓	✓	✓		
Commercial/ retail land use area	✓	✓		✓		✓	✓		✓		
Transit access		✓	✓	✓		✓			✓	✓	✓
… over 200 possible variables											

Example: Street Connectivity

		AUS	BEL	BRA	COL	CZR	DEN	HK1	MEX	NZ	UK	US
G5NIC	GIS, 500M, SN BUFF, INTERSECTION, COUNT	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
G5NID	GIS, 500M, SN BUFF, INTERSECTION, DENSITY	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
G5NIDG	GIS, 500M, SN BUFF, INTERSECTION, DENSITY, SIGNALIZED	N	N	N	N	Y	N	N	N	N	N	N
G5NIDU	GIS, 500M, SN BUFF, INTERSECTION, DENSITY, UNSIGNALIZED	N	N	N	N	Y	N	N	N	N	N	N
G1NIC	GIS, 1KM, SN BUFFER, INTERSECTION, COUNT	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
G1NID	GIS, 1KM, SN BUFFER, INTERSECTION, DENSITY	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
G1NIDG	GIS, 1KM, SN BUFF, INTERSECTION, DENSITY, SIGNALIZED	N	N	N	N	Y	N	N	N	N	N	N
G1NIDU	GIS, 1KM, SN BUFF, INTERSECTION, DENSITY, UNSIGNALIZED	N	N	N	N	Y	N	N	N	N	N	N

Example: Public Transit

		AUS	BEL	BRA	COL	CZR	DEN	HK1	MEX	NZ	UK	US
G1NTDA	GIS, 1KM, SN BUFF, TRANSPORTATION, DENSITY, OVERALL	N	Y	Y	Y	Y	Y	Y	PENDING	Y	Y	Y
G1NTDB	GIS, 1KM, SN BUFF, TRANSPORTATION, DENSITY, BUS ONLY	N	Y	N	N	Y	Y	N	PENDING	Y	Y	Y
G1NTDR	GIS, 1KM, SN BUFF, TRANSPORTATION, DENSITY, RAIL ONLY	N	Y	N	N	Y	Y	N	N	Y	Y	Y
GDXTXA	GIS, DISTANCE (NETWORK) TO NEAREST, TRANSPORTATION, OVERALL	N	Y	Y	Y	Y	Y	Y	PENDING	Y	Y	Y
GDXTXB	GIS, DISTANCE (NETWORK) TO NEAREST, TRANSPORTATION, BUS ONLY	N	Y	N	N	Y	Y	N	PENDING	Y	Y	Y
GDXTXR	GIS, DISTANCE (NETWORK) TO NEAREST, TRANSPORTATION, RAIL ONLY	N	Y	N	N	Y	Y	N	N	Y	Y	Y

Other Resources

- Frank et al. *Br J Sport Med,* 2010

- Twin Cities Walkability
  - Provides step-by-step GIS guide
  - http://designforhealth.net
  - Note: IPEN Templates supersede Forsyth's definitions

- UD4H / IPEN CC Webinars
  - Recorded for on-demand playback
  - In planning stages

IPEN Training Needs: Webinars

From Jim Chapman at UD4H

Webinar Topics	Brazil	Israel	Hong Kong	Denmark	Portugal	Spain	Belgium	Czech Republic
1. Development of Network-Based Buffers		x	x		x	x		x
2. Common Built Environment Measures – Counts, Densities, Crow-fly & Network Distance to Nearest Amenity			x		x	x		
3. Aggregating Parcel Data to Participant Buffers	x	x	x		x	x	x	x
4. Defining Microscale Destinations – Identifying Amenity Destinations, Business Clusters & Places for Physical Activity			x		x	x		
5. Preparing Park Data – Polygons to Park Boundary Points			x		x	x		
6. Creating Walkable & Pedestrian Enhanced Road Networks	x	x	x		x	x	x	
7. Creating Public Transit Variables	x	x	x	x	x	x	x	
Other suggestions?								

Questions

- When are countries planning to start individual-level GIS work?

- When would be a good date to have new Templates completed?

- Recommendation on best way to provide trainings?
  - e.g. email, 1-on-1 visits, groups, phone or GoToMeeting conference calls, webinars, international conferences?