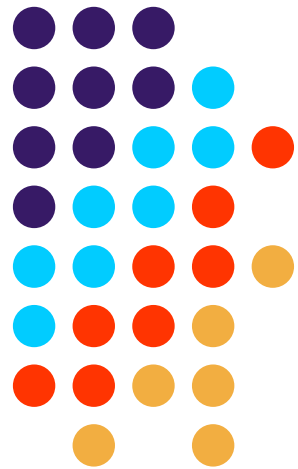
DATA STANDARDS

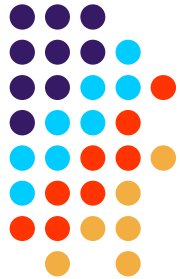
www.TorontoHealthProfiles.ca



About the Data

- Some of the health data on the website has never been prepared for use at the small area level before.
- The researchers, epidemiologist, analysts and geographers preparing the information for the website continue to assess and test new data sets to add variables to the website.
- The geographic focus is relevant to understanding population health and social conditions.

Community Health Planning



The founding partners established the website to provide Toronto communities with information relevant to health planning and decision making with the overall goal of reducing health inequalities. The information is provided at various levels of geography to help organizations with local service areas identify unique needs and changes; and, to help city-wide organizations identify priority areas for strategies and partnerships.

Toronto Health Profiles

Search

www.TorontoHealthProfiles.ca

Home | Partners | Thematic Maps | Health Profiles | Data Holdings | Variable Definitions | Resources | Contacts | Feedback | Help

Home

Toronto Community Health Profiles Partnership

Printable Version

St. Michael's Hospital — Leading with Innovation, Serving with Compassion

SETo

Toronto District Health Council / Le Conseil régional de santé de Toronto — www.tdhc.org

Toronto Public Health

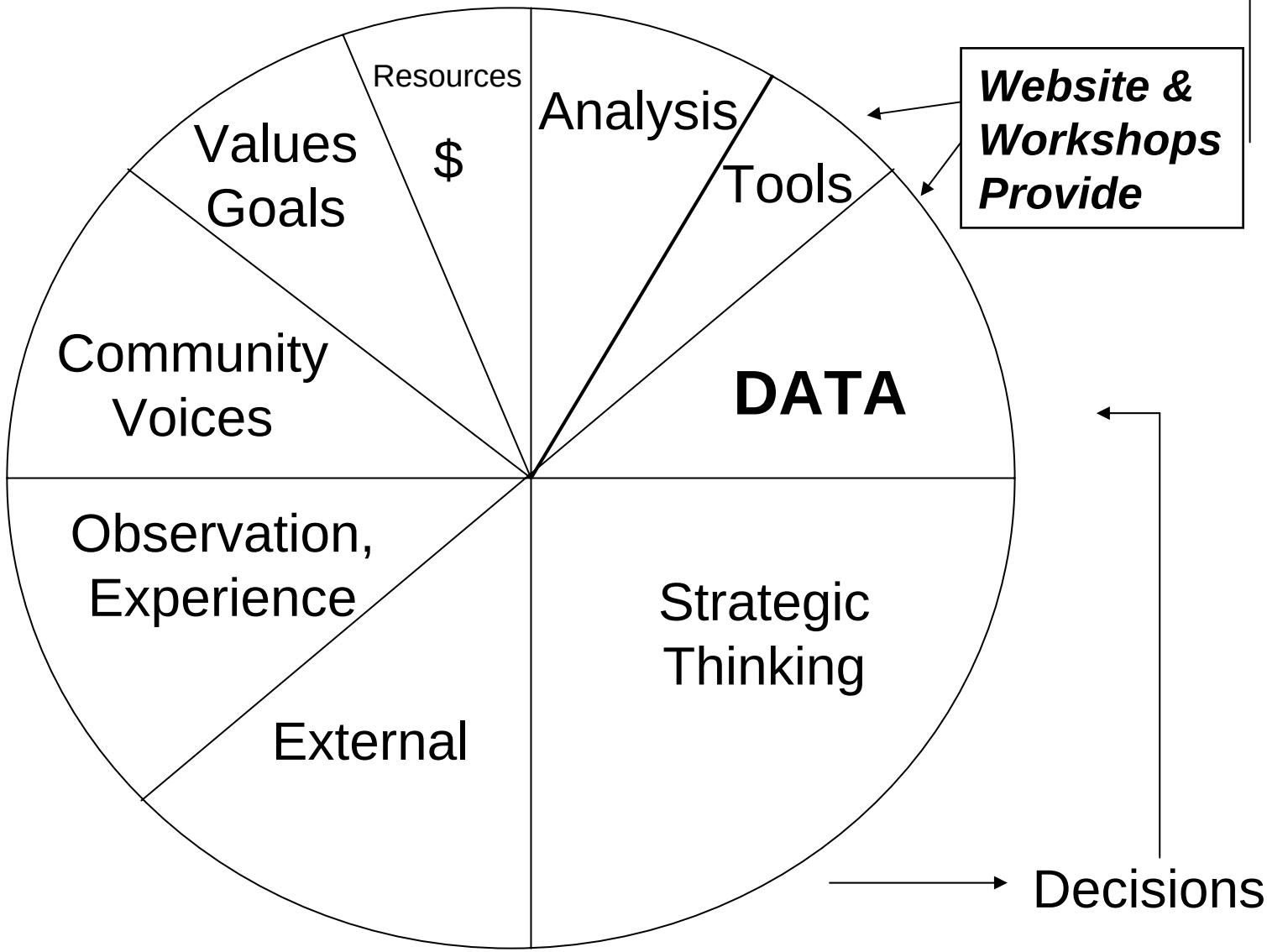
Wellesley Central — Promoting Urban Health

A partnership designed to facilitate access to information for health planning with the overall goal of producing action to reduce health inequalities.

Using the Data

- The data on the website is meant to be used along with other information (see planning circle on the next slide).
- A series of workshops, user guides, and resources are being provided to help users understand, interpret and apply the information with an awareness of its limitations.

Community Health Planning



Website & Workshops Provide

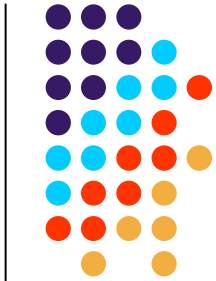
Decisions

Patychuk, March 2005

Purpose of the Data Standards

This guide describes the steps taken to ensure that the information on the www.TorontoHealthProfiles.ca website is accurate, complete and useful and that users are aware of the limitations. Since users are interested in looking for differences between areas, the objective is to reduce the amount of difference that may be due to the quality of the data (variability, small numbers, small sample size, calculation errors, representativeness of the sample, misunderstanding or misinterpretation of the meaning of the indicator, etc). Epidemiological practice standards and small area analysis guidelines used by other organizations were consulted in developing these data standards.

Limitations Not Addressed



- The data may be used in research to identify the range of possible reasons for observed differences in health, but the maps and tables don’t do this on their own. Caution is advised in drawing conclusions based on limited data.
- The geographic focus of the website is less relevant to understanding the health of communities that are not geographically concentrated (such as specific ethnic communities, people who are homeless, recent immigrants, etc.).

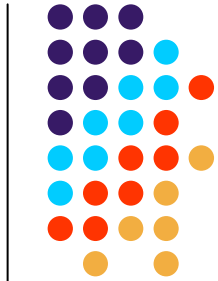
Data Standards: Outline

- Selecting Indicators
- Health Indicators
- Spatial Issues
- Demographic Effects
- Random Effects
- Reporting Standards
- Quality Assurance (QA)
- Data Sources & Limitations

Selecting Indicators: Criteria

Indicators are Relevant to Reducing health Inequalities	Comparable across time and place; Common or “core” indicator (APHEO, CPHI, DHC, ICES, FCM, Statistics Canada, Other cities, CTPHC); Address complexity, size and diversity of Toronto; Indicator has demonstrated relationship to health inequalities or is being used in a similar context.
Indicators are of Good Quality	Consistent definition and data collection methods are used, data integrity is maintained, data is obtained from reliable sources, calculation is transparent and can be reproduced, missing cases are identifiable, standards for small area analysis and reporting are applied, data is available for all areas of the city, limitations on use and interpretation are reported.
Supplements existing sources of information	Indicators supplement other community profile initiatives and are developed in coordination with other information providers.
Useful	User reviewed, user input, user-identified priorities.

Health Indicators



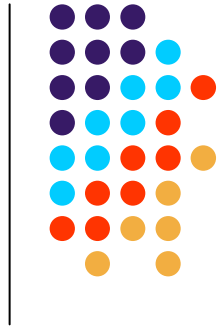
- Age specific or age standardized rates
- Indicator definitions
- Health indicators across the life span will be included: e.g. determinants, behaviours, perceived health, use of prevention and treatment, health outcomes, mortality, disease prevalence, medications, etc.
- Confidence Intervals (C.I.) are calculated to identify rates higher (H), lower (L) or not significantly different (NS) from city rate 19 times out of 20.
- Rate Ratios: area rate divided by city rate to identify policy significance – size of health gap (e.g. 1.2 times > city rate)

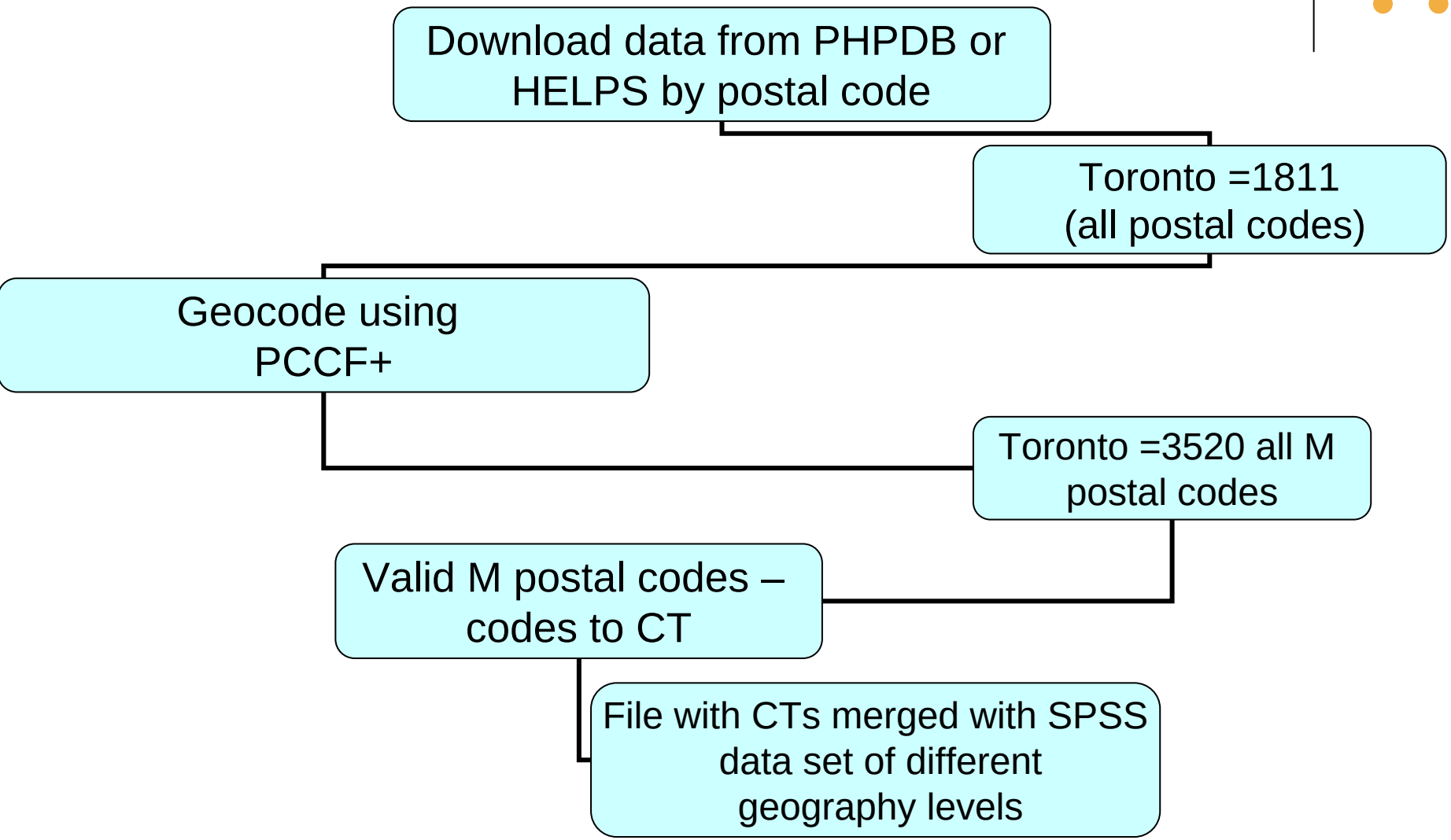
Data is based on Place of Residence

- All health data is based on residence* of individuals not where service was provided.
- All health data is geocoded to census tracts which are aggregated up to the other geographic levels.
- The total in the profiles is the aggregate of all geocoded data (excludes data without a valid Toronto postal code). So the total on the profiles may be up to 2-4% less than the city total reported elsewhere for city data that are not based on geocoding.

*The only data based on place of occurrence is police data and mapping of service sites

Geocoding

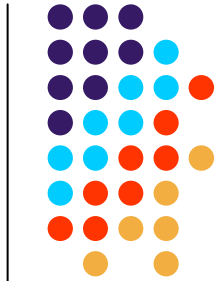



Download data from PHPDB or HELPS by postal code
Toronto =1811 (all postal codes)
Geocode using PCCF+
Toronto =3520 all M postal codes
Valid M postal codes – codes to CT
File with CTs merged with SPSS data set of different geography levels

Area versus Individual Measures

- Neighbourhoods and planning area rates represent an “average” of the individuals living in the area. Individual, family and households incomes can vary widely as many Toronto areas are mixed income communities.
- Area rates cannot be assumed to apply to all the individuals living in the area. For example if 40% of a neighbourhood’s residents are low income, and 40% of residents report using a health care service, it cannot be assumed that all those using the service were the low income residents
- Cannot attribute SES characteristic to an individual based on area rates

Accounting for Demographic Effects



Demographic Composition

- Variations based on the age and gender make up of an area can explain the observed differences in health events that are known to vary by age and gender.
- Example: a neighbourhood with a high proportion of older adults 75+ will have higher rates of chronic diseases and disabilities that may be explained by these age differences

Strategy for accounting for age/gender effects

- Age standardized rates by gender
- Age specific rates where the events or indicators are concentrated (e.g. mammograms among females age 50-69)
- Identify sites located in an area that include a concentration of specific populations (e.g. residences for pregnant teens, long term care facilities. etc.)

Minimizing Random Effects

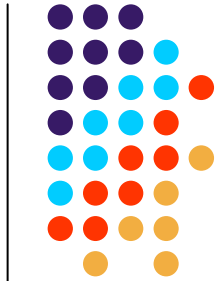
Random noise:

- Variations based on size of numerator and denominator that can lead to instability in rates because the event is infrequent (rare events) or the number of people in the area that the rate applies to is small.
- Example: A small increase in the number of births among a small population of female teens could double the rate but it reflects too small a number of events to be important for planning. It could be a one-time thing.

Strategy:

- Reporting standards
- Combine up to 5 years to obtain reportable information
- Combine geographic areas – report only for larger areas
- Coefficient of Variation used in CCHS survey data
- Confidence intervals

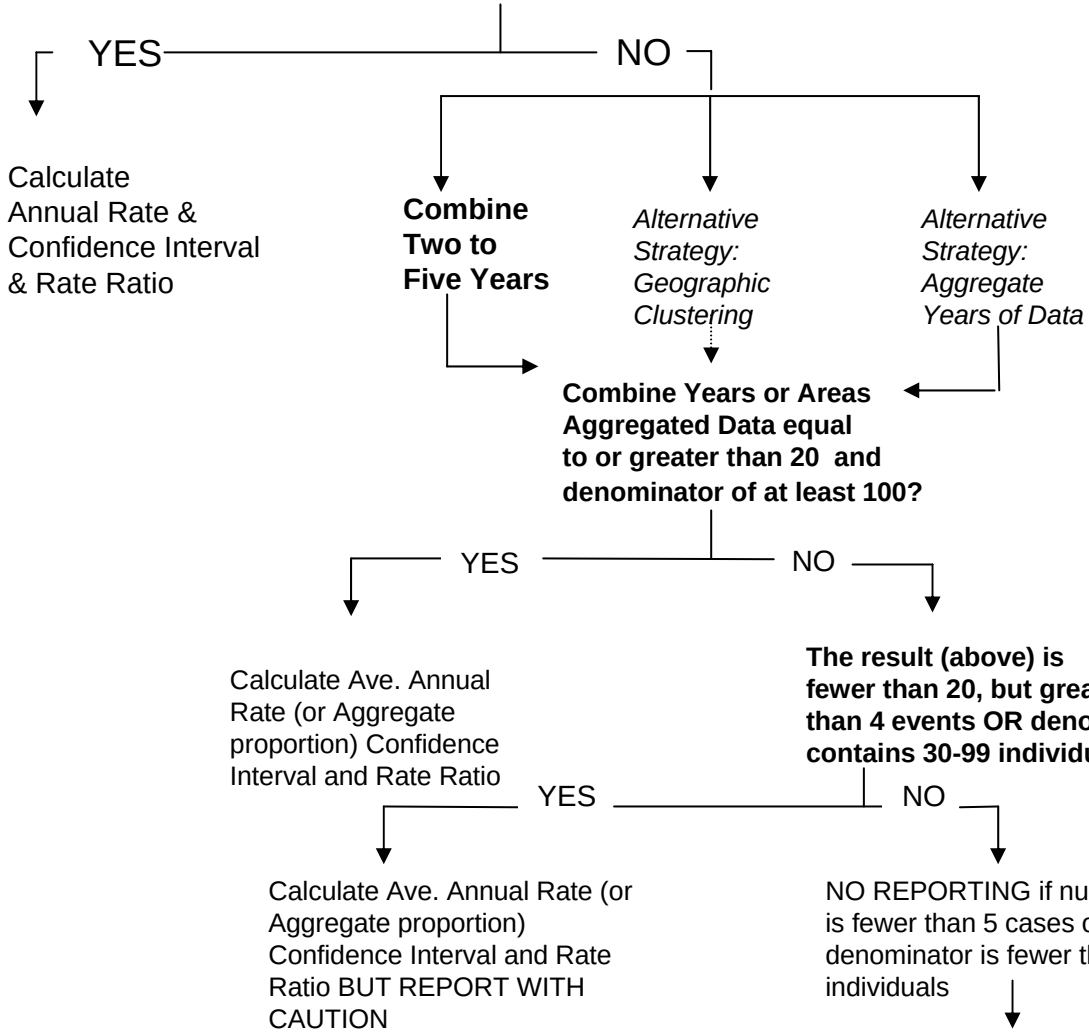
Ethics & Reporting Standards



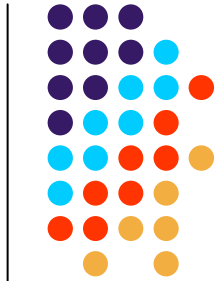
- Full reporting if numerator at least 20 and denominator at least 100
- Reporting with caution if numerator contains 5-19 events OR denominator contains 30-99 individuals
- No reporting if numerator less than 5 individuals or denominator fewer than 30
- Aggregate data for areas or years (2-5 years) for larger sample or population
- No individual level data

TCHPP SMALL NUMBERS FLOW CHART

Annual # of cases equal to/greater than 20 and denominator of at least 100?



Mapping Standards



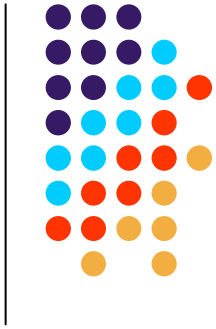
- Map variable at the smallest geographic level for which the majority of extreme, policy relevant rates (e.g. 20% > or < than total rate) are statistically significant (95% confidence intervals)
- Data must be reportable for at least 80% of the units in the geographic level (e.g. if rates for one of the minor areas cannot be reported, the variable will not be mapped at that geographic level).
- For Health Indicators, identify which rates are statistically significant

CCHS Reporting Standards

Use of the Canadian Community Health Survey (CCHS) data requires:

- 1. Checking the unweighted estimates to make sure that the numerator of each cell is not less than 10 for the Ontario Share File, or 30 for the PUMF.
- 2. Checking the coefficient of variation (either using CV look-up tables or bootstrapping to create CVs) and follow the release guidelines

Apply CCHS Sampling Variability Standards



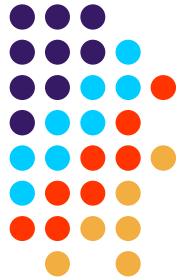
Unqualified (CV= 0.0 – 16.5) Estimates general unrestricted release.

Marginal (CV=16.6 – 33.3) Estimates considered for general unrestricted release but should be accompanied by a warning of high variability associated with estimates. (Footnote on table)

Unacceptable (CV> 33.3) Estimates of unacceptable quality. Conclusions based on these data will be unreliable and most likely invalid and should not be reported.

The CCHS 1.1 data used on the website was prepared for this purpose by Statistics Canada.

Data Quality (QA)

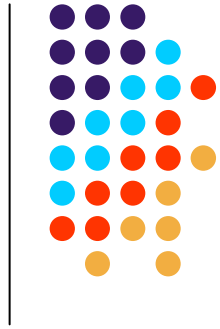


Data Checks:

- Consistent with published data
- Consistent with internal reports/analysis
- Confirmed by independent analysis
- Confirmed by rerunning program
- Do manual computations
- Incorporate formula checks to worksheets
- State data limitations, % missing, representativeness of sample
- Documentation of QA checks

Data Sources & Limitations

Canadian Community Health Survey (CCHS) 1.1 2000/01



Strengths:

- Detailed information on individuals (e.g. income, education, ethnicity)
- 1st person accounts of health system experiences and health status (administrative databases only describe utilization)
- Useful as a relative measure of the range of differences

Limitations:

- Small sample size (2382) – no respondents in some neighbourhoods & need to aggregate to large geography – wide confidence intervals
- May not be representative of entire population in areas
- Crude indicators – not age standardized or age specific
- People under-report certain conditions (eg. Chronic conditions) and socially undesirable behaviour (eg. Smoking during pregnancy) leading to underestimates of prevalence
- People over-estimate socially desirable behaviours (eg. Exercise, fruit & vegetable consumption)

CCHS Representativeness

Assessment based on % age 15+

- in owner households,
- immigrants,
- age 65+
- female/male

CCHS 1.1 weighted sample was compared to the 2001 Census 15+ in Households for rate differences >15%; percentage point differences >10%; change in ranking out of 15; and, change in High/Low/Similar clustering. In the majority of cases there was little change in the relative ranking of the 15 areas. Therefore the 15 Minor Health Planning Areas are potentially useful for demonstrating the range of health differences. Their usefulness will be improved by combining several survey years (1.1 with 2.1 and 3.1) to increase sample size and better assess representativeness and significance.

CCHS Representativeness: 15 Minor Health Planning areas (MHPAs)

CCHS 1.1 sample is comparable to area populations on the variables assessed for 5 MHPAs: 2b-York Weston, 3c-Midtown, 4a-Danforh East York; 4b-Downtown & Waterfront and 5c-Scarborough Centre
CCHS 1.1 sample underrepresents seniors but is comparable for other variables assessed for 3 MHPAs: 1a-Rexdale, 5a-Agincourt & 5b-Rouge
CCHS 1.1 sample underrepresents homeowners but is comparable for other variables assessed for 2MHPAs:1b-Etobicoke South & 2a-Humber Downsview
The CCHS 1.1 sample for 2 MPHAs (3a-Willowdale and 3b-North Toronto Don Mills) over-represents seniors (thus more females in 3b) but is comparable on the other variables assessed.
The CCHS 1.1 sample for 2c York Junction underrepresents seniors and owners; 2d-Parkdale underrepresents seniors and immigrants, and 5d-Scarborough Cliffs underrepresents seniors and owners and over-represents immigrants (3 MHPAs). The CCHS is less useful for these areas.

Data Sources & Limitations (cont'd)

Canada Census 1991,1996, 2001
(Statistics Canada)

Strengths:

- Best (and only) source of social and demographic info for the entire population (some exceptions)
- Large # of variables : over 1,500

Limitations:

- Census undercount: 5.17% for the Toronto Census Metropolitan Area (CMA); underrepresented groups
- Data suppression, particularly at DA level
- Census tracts only in urban areas limit comparisons
- Only every 5 years

Data Sources & Limitations (cont'd)

Physician Claims (OHIP)

Strengths:

- Can answer: "Who is using services and what kind?"
- Only comprehensive source of population health coverage & provision of publicly-paid health services
- Laboratory and radiology claims include CHCs

Limitations:

- Excludes CHCs for physician visits (e.g. diabetes)
- Health insurance addresses out-of-date
- No individual level socioeconomic or cultural info available

Data Sources & Limitations (cont'd)

Vital Statistics - (births and deaths)

Live Birth Database (PHPDB), Health Planning System (HELPS), MOHLTC

Strengths:

- Includes country of birth (HELPS only)
- Links baby to mother for analysis of singleton LBW by age, parity, pregnancy type (HELPS only)

Limitations:

- missing unregistered births
- missing postal codes (potentially over 3%)
- 2- 3 yr time lag in data availability

Data Sources & Limitations (cont'd)

Hospital Inpatient Data - Canadian Institute for Health Information (CIHI); PHPDB

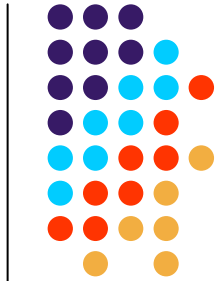
Strengths:

- Up-to-date postal codes
- Current – I year time lag

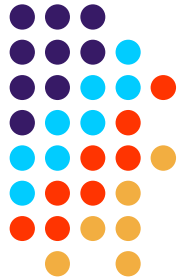
Limitations

- No mental health data available
- Excludes out of hospital births
- Missing postal codes approximately 2%
- No SES or ethnicity info available

Data Issues



- Balancing making the information user-friendly with the providing detail needed for accurate and appropriate use and understanding of the information
- Sustainability, capacity to update data in the future
- Reducing the resources required for data conversion through developing a user-driven interactive site
- Responding to potential health inequalities that are identified on the site
- Comprehensiveness across the range and breadth of health planning needs



THANK YOU!

VISIT the “Resources” TAB for more Information, and the “About the Data” TAB for Variable Definitions.