

Data Collection & Analysis for Disease Accounts

Dr Ravi P. Rannan-Eliya

Estimating Expenditure by Disease Training Workshop

WHO, Geneva

August 20-24, 2007



Outline

- General strategy
 - Identification of spending areas
 - Review of data sources
- Inpatient spending
- Outpatient spending
- Pharmaceuticals

Reminder

Goals of a Disease-Specific Account (DSA)

- To classify all disease-relevant expenditures by specific diseases
 - Does not include all health expenditures
- Includes:
 - Primarily - Treatment of individual patients
 - HC1 – HC5

ICD-10 Chapters

Chapter	Blocks	Title
<u>I</u>	<u>A00-B99</u>	Certain infectious and parasitic diseases
<u>II</u>	<u>C00-D48</u>	Neoplasms
<u>III</u>	<u>D50-D89</u>	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
<u>IV</u>	<u>E00-E90</u>	Endocrine, nutritional and metabolic diseases
<u>V</u>	<u>F00-F99</u>	Mental and behavioural disorders
<u>VI</u>	<u>G00-G99</u>	Diseases of the nervous system
<u>VII</u>	<u>H00-H59</u>	Diseases of the eye and adnexa
<u>VIII</u>	<u>H60-H95</u>	Diseases of the ear and mastoid process
<u>IX</u>	<u>I00-I99</u>	Diseases of the circulatory system
<u>X</u>	<u>J00-J99</u>	Diseases of the respiratory system
<u>XI</u>	<u>K00-K93</u>	Diseases of the digestive system
<u>XII</u>	<u>L00-L99</u>	Diseases of the skin and subcutaneous tissue
<u>XIII</u>	<u>M00-M99</u>	Diseases of the musculoskeletal system and connective tissue
<u>XIV</u>	<u>N00-N99</u>	Diseases of the genitourinary system
<u>XV</u>	<u>O00-O99</u>	Pregnancy, childbirth and the puerperium
<u>XVI</u>	<u>P00-P96</u>	Certain conditions originating in the perinatal period
<u>XVII</u>	<u>Q00-Q99</u>	Congenital malformations, deformations and chromosomal abnormalities
<u>XVIII</u>	<u>R00-R99</u>	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified
<u>XIX</u>	<u>S00-T98</u>	Injury, poisoning and certain other consequences of external causes
<u>XX</u>	<u>V01-Y98</u>	External causes of morbidity and mortality
<u>XXI</u>	<u>Z00-Z99</u>	Factors influencing health status and contact with health services
<u>XXII</u>	<u>U00-U99</u>	Codes for special purposes

ICD-10 Chapter detail

Neoplasms

(C00-D48)

C00-C97 Malignant neoplasms

C00-C75 Malignant neoplasms, stated or presumed to be primary, of specified sites, except of lymphoid, haematopoietic and related tissue

C00-C14 Lip, oral cavity and pharynx

C15-C26 Digestive organs

C30-C39 Respiratory and intrathoracic organs

C40-C41 Bone and articular cartilage

C43-C44 Skin

C45-C49 Mesothelial and soft tissue

C50 Breast

C51-C58 Female genital organs

C60-C63 Male genital organs

C64-C68 Urinary tract

C69-C72 Eye, brain and other parts of central nervous system

C73-C75 Thyroid and other endocrine glands

C76-C80 Malignant neoplasms of ill-defined, secondary and unspecified sites

C81-C96 Malignant neoplasms, stated or presumed to be primary, of lymphoid, haematopoietic and related tissue

C97 Malignant neoplasms of independent (primary) multiple sites

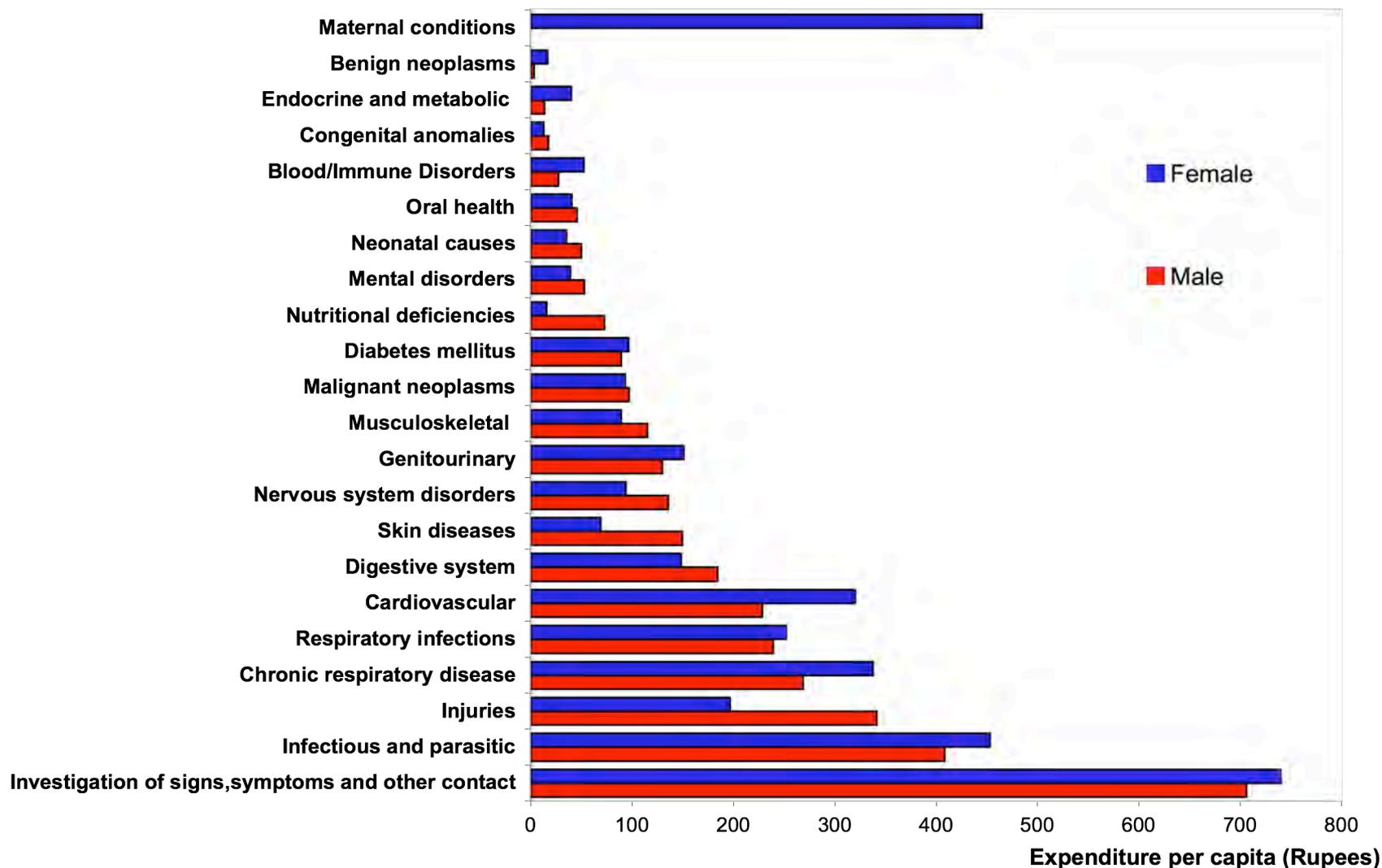
D00-D09 In situ neoplasms

D10-D36 Benign neoplasms

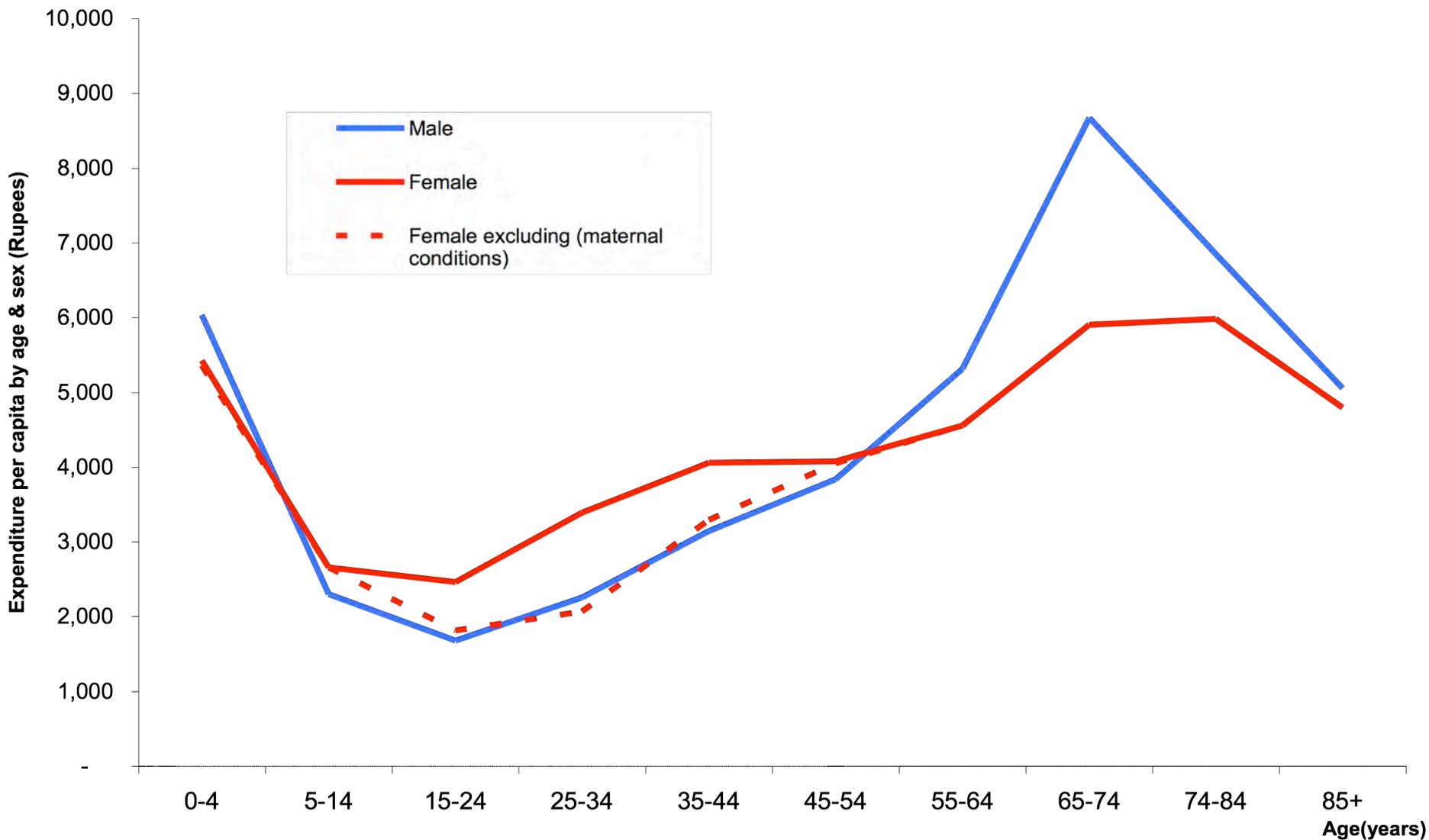
D37-D48 Neoplasms of uncertain or unknown behaviour [see note before D37]

Health expenditures by condition

Sri Lanka 2005

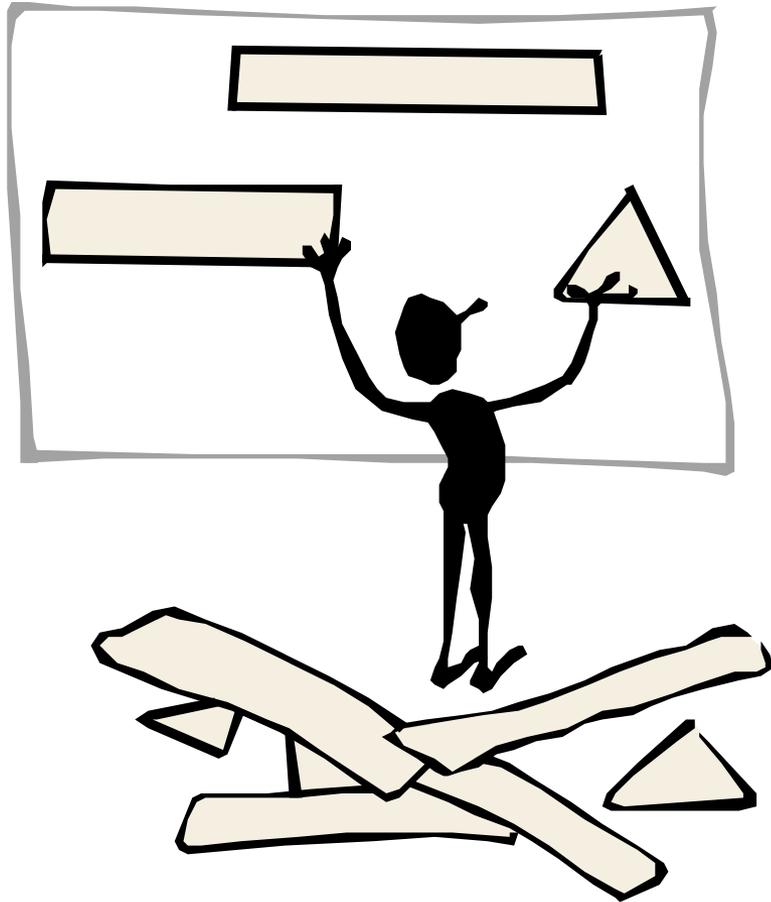


Health expenditures per capita by age and sex, Sri Lanka 2005



General Strategy

General Strategy



- Review available data sources and possible methods
- Identify discrete components of spending for analysis
- Primary/secondary data collection & coding
- Data analysis

Australia: Spending Areas



- Hospital - inpatients
- Hospital - outpatients
- Aged care homes
- Non-hospital medical services
- Pharmaceuticals
- Other professional services
- Dental services
- Research

Approaches for general patient treatment

1. Use existing databases of financial transactions coded by ICD-10
 - Health insurance records
 - Patient-based payment systems in integrated systems
2. Estimate distribution of expenditures using cost distribution keys
 - E.g., Patient activity numbers
 - E.g., Input distributions
3. $P \times Q$ approach - Do not recommend for a full disease analysis

Cost distribution keys

Basic method

$$\text{Disease A expenditures} = \text{Total expenditures} \times \frac{\text{No. of patients with Disease A}}{\text{No. of all patients}}$$

Keys

- Patient outpatient visits, admissions
- Bed-days, operations, lab tests, prescriptions
- Drug costs, staff costs, lab costs
- Doctors' time, nursing staff time, other staff time
- Utilities costs, administrative costs

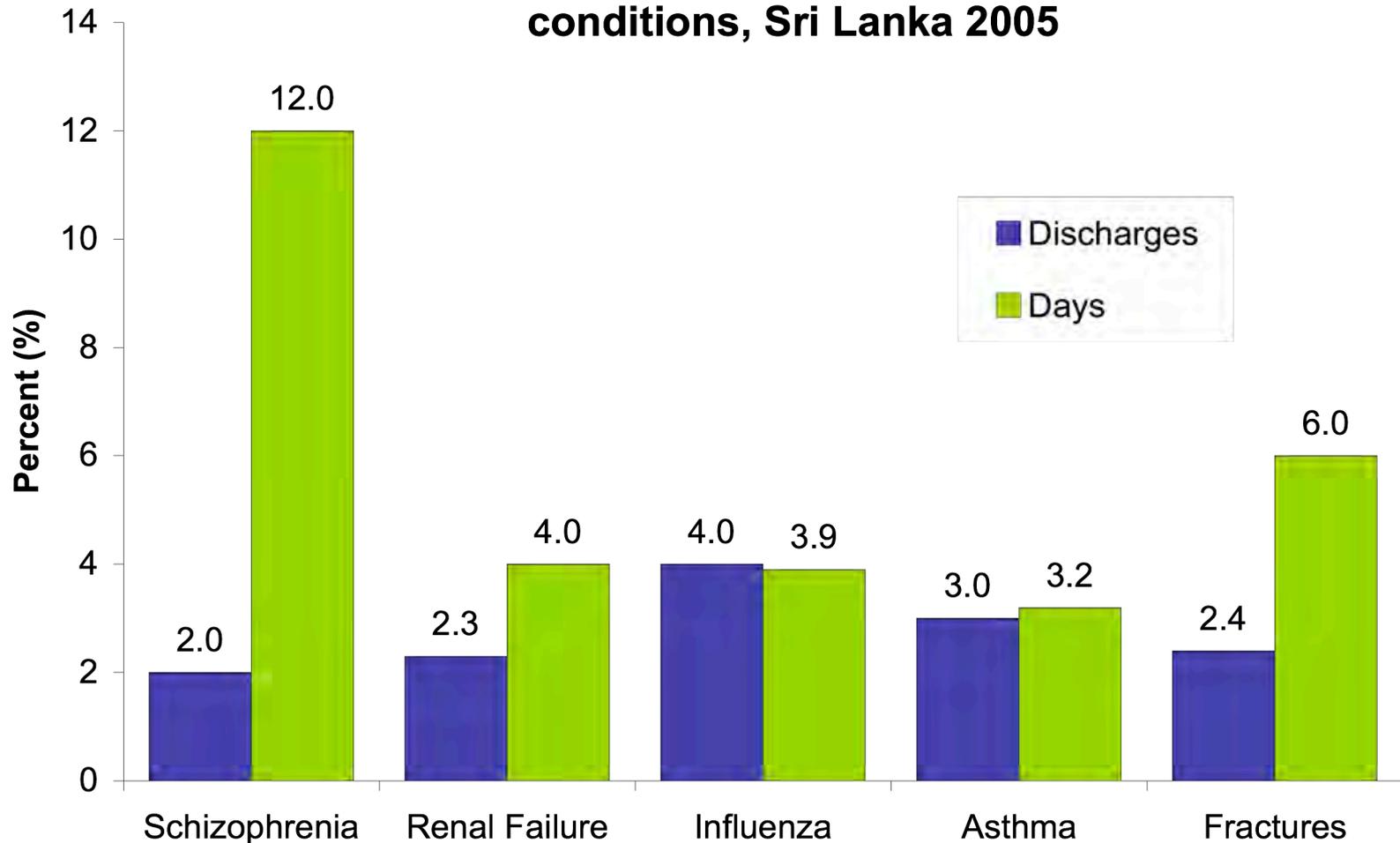
Inpatient spending

Public hospital inpatient

- Basis of measurement is **cost** (not price)
 - Both budgetary spending and user charges should be distributed according to cost of production
 - In countries where hospital budgets do not differentiate inpatient from outpatient inputs, estimate of inpatient share should be based on cost surveys
- Starting point should be number of admissions by disease/age/sex
 - More reliable estimates will refine cost drivers

Why admission numbers can be misleading

Percentage of discharges and bed-days, selected conditions, Sri Lanka 2005



Secondary data sources

- Secondary data sources
 - Public sector data on inpatient admissions
 - Household health surveys
 - Unlikely to be reliable or samples to be large enough
- Potential problems
 - Aggregate data with no cross-tabulations possible between disease and age/sex
 - Diseases are reported in aggregate categories - not ICD-10
 - Disease tabulations are for top most common conditions, but not for all conditions
 - No stratification available by hospital type

Primary
data
collection

National Inpatient Discharge Survey Design

- Nationally representative
 - Different epidemiological regions, hospital types
- Retrospective or prospective?
 - Depends on inpatient records system - If discharge records are generated, retrospective is better using year-wide sample
- Better to link to hospital cost data/survey

National Inpatient Discharge Survey Design

- Sampling
 - N ~ 5,000 - 10,000 minimum
 - Facility sample stratified by type and region
 - Sample patients within facilities - systematic or random
 - Use HIS data to generate sampling weights or collect turnover data by facility
- Record collection
 - Extract data on-site
 - ICD-10 diagnosis or diagnosis in original words
 - (Second diagnosis?)
 - Age, sex
 - Other data: LOS, inputs (operations, drugs, lab tests, etc)

National Inpatient Discharge Survey Design

- Data entry and processing
 - ICD-10 coding in office
 - Use data entry software to minimize data entry errors
 - Error checking using post-survey validation analyses
- Cost analysis
 - Use programmable statistical software (Stata, SPSS)
 - Ideally link patient survey to hospital cost survey which uses step-down cost analysis to identify cost centres
 - Distribute cost centre costs using inputs
- Examples
 - Ward costs - bed-days
 - IP drugs - actual drug costs/numbers
 - Theatre costs - weighted sum of operations
 - Radiology - weighted sum of X-rays

National Inpatient Discharge Survey Design

- Advanced tips
 - Redistribute missings using known distribution of relevant cases
 - Eg: distribute missing age in lung cancer cases using age distribution of all lung cancers
 - For low-frequency diseases where better estimates of total admission numbers are available, apply post-stratification weights, e.g., TB, malaria, HIV
 - If aggregate admission data are available by larger disease categories, can use post-stratification weights to adjust sample to match larger categories

Private hospital inpatient

- Basis of measurement is **price** (not cost)
 - Obtaining revenue data by ICD-10 may be difficult
 - Alternative is to survey admission numbers
- Alternatives
 - Household survey data might be used to generate age-sex or major disease category distributions
 - Insurance claims data might provide information - caution if insurance population is not representative

Outpatient spending

Outpatients

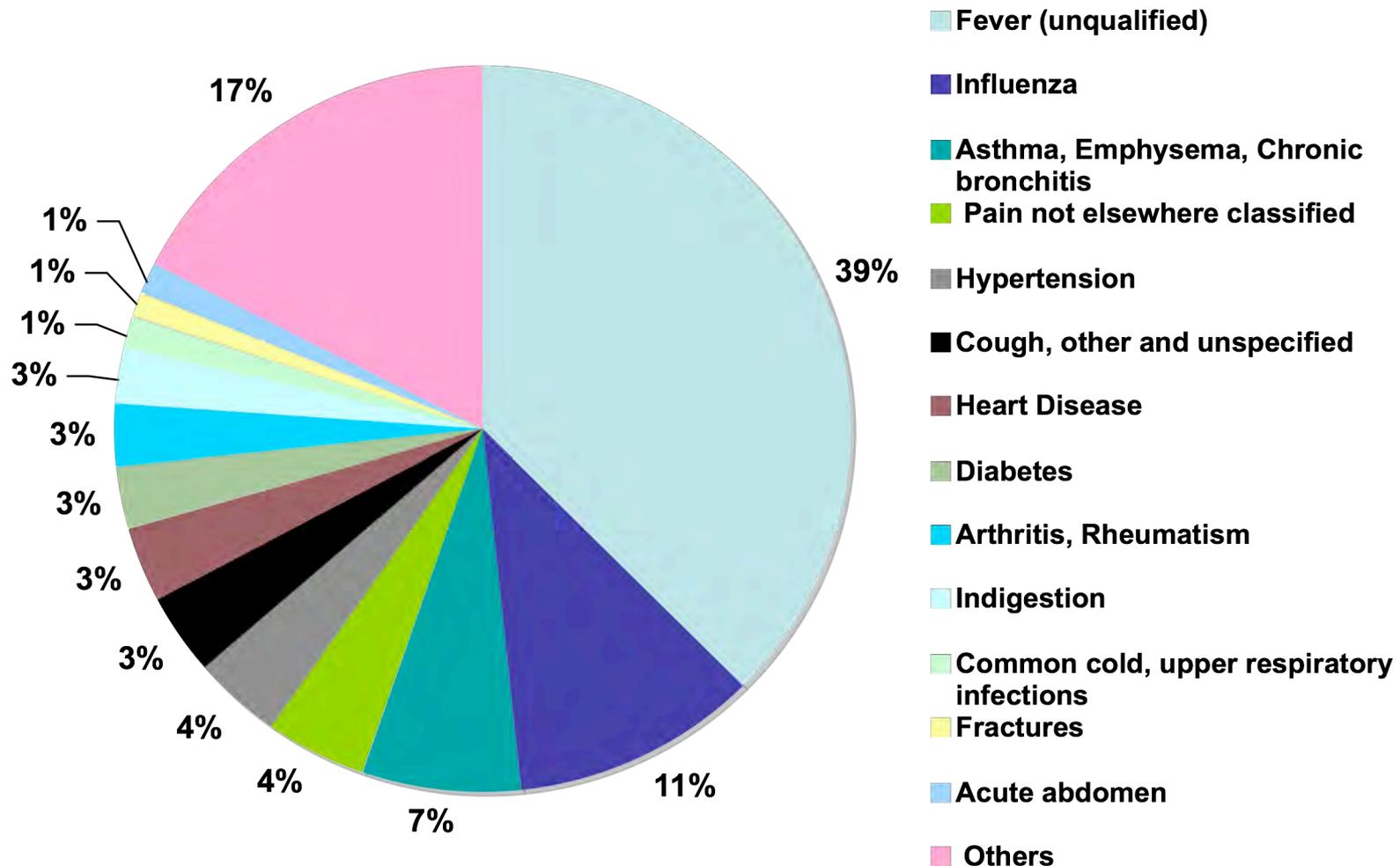
- Basis of measurement is **cost** (public sector) or **price** (private sector)
 - Both budgetary spending and user charges should be distributed according to cost of production
 - In countries where hospital budgets do not differentiate inpatient from outpatient inputs, estimate of outpatient share should be based on cost surveys
- Starting point should be number of outpatient visits by disease/age/sex
 - More reliable estimates will refine cost drivers

Secondary data sources

- Secondary data sources
 - Public sector HIS data on outpatient visits/attendances
 - Public or Private provider patient surveys
 - Household health surveys
 - Maybe OK for age and sex
- Potential problems
 - Diseases are reported in aggregate categories - not ICD-10
 - Disease tabulations not for all conditions
 - Provider surveys may not cover all private provider types
 - Household surveys cannot report ICD-10 diagnoses, but usually a mix of symptoms and diseases

Primary
data
collection

Household survey data on OP disease, Sri Lanka 2004



Potential approaches

1. Patient survey at providers

- Public sector - may work
- Private sector - very difficult
 - How to sample? and how to weight patients?
 - Private sector cooperation?
 - Traditional providers?
 - Pharmacies, shops?

2. Combination of household survey and patient survey

Combining household and provider surveys

Household survey

- Existing health survey or new survey
 - Visit numbers
 - Expenditures
- Distribution by provider types - *may be reliable*
- Distribution of age, sex - *may be reliable*
- Distribution of problems in patients' own words - *likely to be unreliable*

Patient survey

- Sample survey that records patients' problems in own words & diagnosis (ICD-10)
- Provides:
 - Mapping of problems to ICD-10 codes
 - Distribution of ICD-10 codes by patient problems

What is ICPC-2?

- “International Classification of Primary Care”
- WHO classification for primary care recording
- Coding system for reason for encounter (not disease)
 - Then code to actual diagnosis <> ICD-10
- References:
 - <http://www.who.int/classifications/icd/adaptations/icpc2>
 - <http://www.globalfamilydoctor.com/wicc/>
 - Rannan-Eliya, Ravi P., Prashanthi Jayawardhane, and Leela Karunaratne. 2003. *Private primary care practitioners in Sri Lanka*. In *Health Policy Research in Asia: Guiding Reforms and Building Capacity*, edited by A. S. Yazbeck and D. H. Peters. Washington, DC, USA: World Bank.

Designing a patient survey

- Public or private or both?
- Sampling?
- Observational or provider-recorded?
- Record
 - Patient problems using ICPC-2 “Reason for encounter” codes
 - Presumed diagnosis (from provider) using ICPC problem codes (maps to ICD-10)
 - Age, sex
 - Medicines prescribed

Example: Mapping of fever in Sri Lanka patient survey

ICD-10 code/condition	% share
R50 Fever	50
B59 Viral disease, other	44
B00.82 Mouth disease, etc	0.6
B08.5 Upp. resp infection, acute	0.6

Pharmaceutical spending

Secondary data sources

- Secondary data sources
 - Public sector data on drugs prescribed by outpatient diagnosis
 - Surveys of pharmacy customers
 - Household health surveys
 - Private expenditure on medicines by diagnosis or problem
- Potential problems
 - Household survey data on medicine expenditures may not be by episode of illness or by individual
 - Diseases are according to self-reported patient problems, not ICD-10

Primary
data
collection

Potential approaches

1. Household expenditure survey available and usable
 - Use same methods as for outpatient analysis to map self-reported problems to ICD-10
2. Use outpatient survey to obtain distribution of each medicine by age, sex and ICD-10 diagnosis of patient
 - Then map sales of each medicine
 - Map at level of WHO-ATC class

Data sources for drug sales

- Industry data on retail sales
 - E.g., IMS-Health data on sales by generic product/WHO-ATC class
- Estimation based on analysis of imports, exports and manufacturing data
 - In practice very difficult to do