

Sources of aggregated registry-based cancer data

Sirpa Heinävaara
Senior researcher, D.Soc.Sc.
Finnish Cancer Registry /
Mass Screening Registry



Aims of the presentation

You will learn that

- Cancer is a large group of diseases
- What kind of cancer registry data are publicly available and where you can find them
- Where you can find long-term data on the occurrence of cancer and what you should be aware of when using these data
- Cancer predictions

Cancer

- A group of over 2000 various diseases
- A rare disease
- Defined by (International Classification of Diseases for Oncology)
 - Primary site/Morphology (**CXX.X**)
 - Type of tumour/Cell type/Histology (**MXXXX/X**)
 - Behaviour/Malignancy
 - Laterality
 - Stage
- Cannot report statistics for each rare disease separately
—→ Reporting malignant tumour by the primary site or combination of them

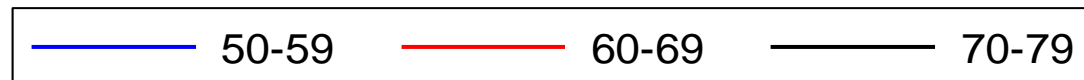
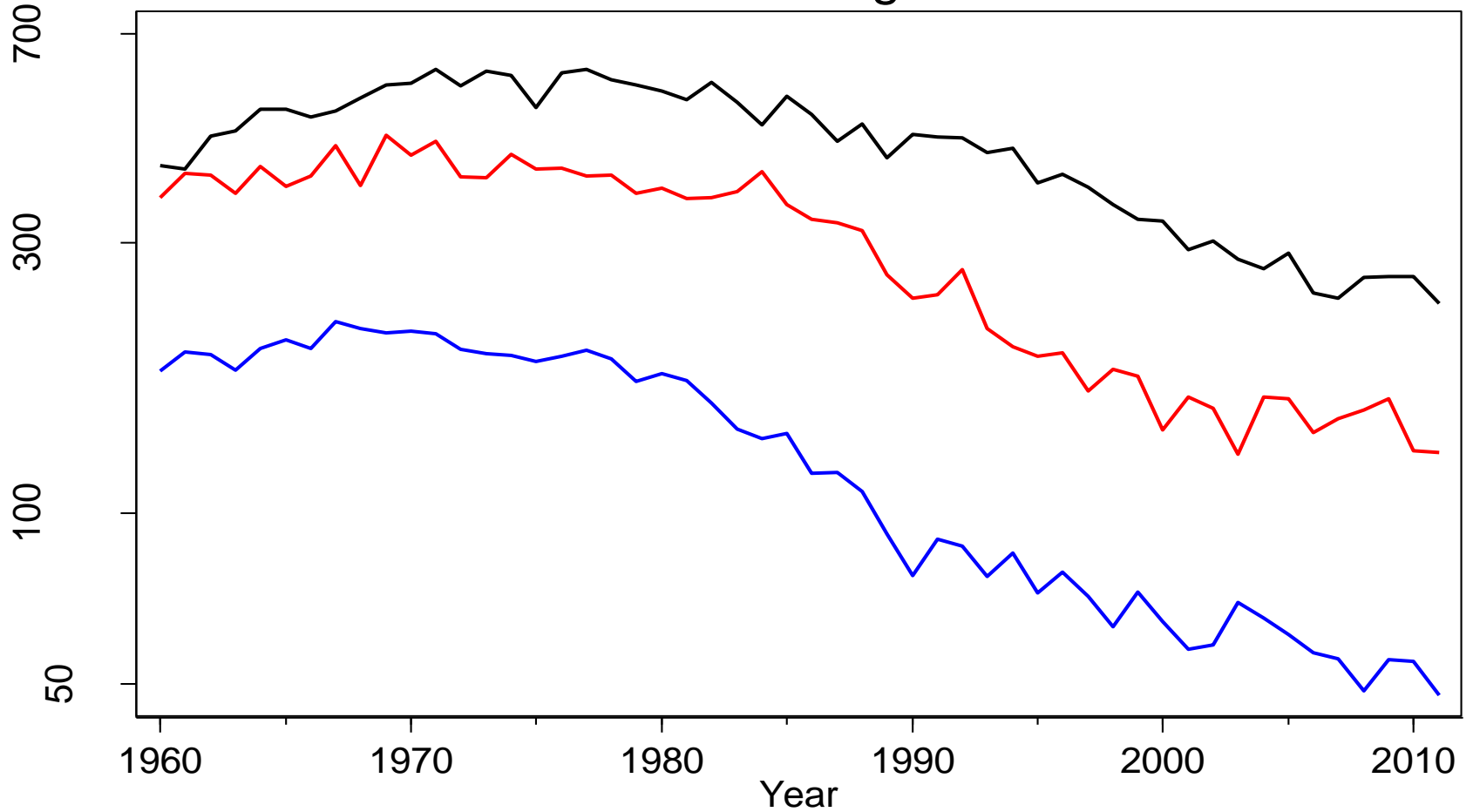
Common measures of occurrence

ABSOLUTE MEASURE: A NUMBER (n)	RELATIVE MEASURE: RATE IN A POPULATION
A number of new diagnoses (cases)	Incidence
A number of cases living with cancer (prevalence)	Prevalence
A number of cancer deaths	Mortality

- Rates are usually calculated by sex and (5-year) age group
- Reliability of numbers and rates are dependent on the coverage of cancer registration, follow-up of individuals, existence of vital statistics, facilities, ...
- For mortality, cause of death should be known

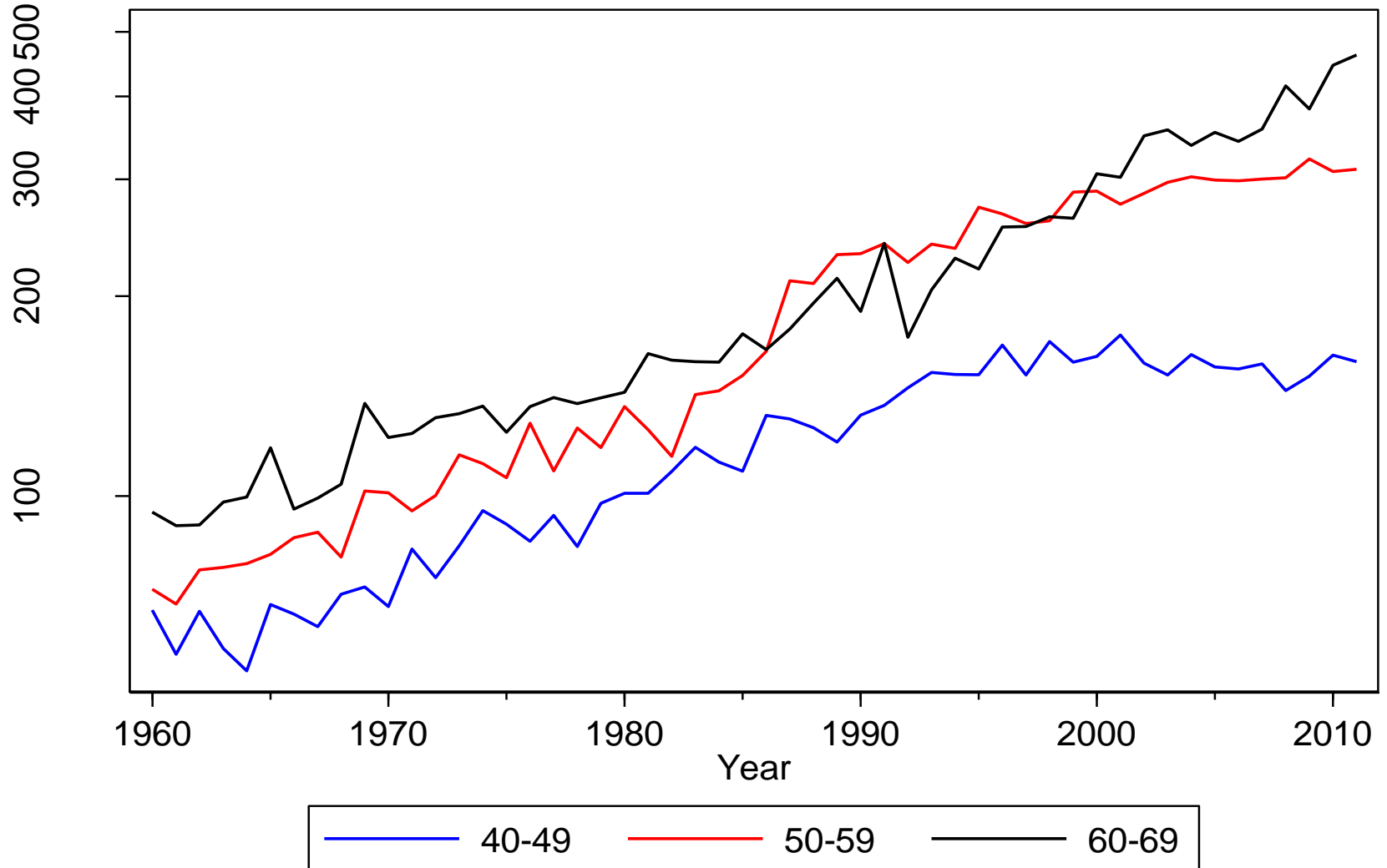
MALES

Incidence of lung cancer



FEMALES

Incidence of breast carcinoma



On common measures of occurrence

Occurrence of cancer is affected by

- Age
- Sex
- SES
- Area
- Ethnicity
- Mass screening programs
- Risk factors and their prevalence (say, smoking)

Information on the most prevalent and established risk factors can be found in **Food, Nutrition, Physical Activity, and the Prevention of cancer**, for example

http://www.dietandcancerreport.org/cancer_resource_center/downloads/Second_Expert_Report_full.pdf

On common measures of occurrence

- Population structure affects both absolute numbers and relative measures
 - Use age-standardized rates to compare measures between different populations or time periods. Different standard population weights are available for standardization (World, European, Nordic).
 - Compare as similarly calculated rates as possible

Sources of information

- GLOBOCAN 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012
<http://globocan.iarc.fr/Default.aspx>
 - 34% (I) and 44% (M) of countries (n=184) had no data at all
 - Several different methods used in estimating the measures
 - ... estimates may not be truly comparable overtime...
 - Online analyses with graphs, tables and maps

Sources of information

- Surveillance, Epidemiology, and End Results programs (SEER) in US: Incidence, prevalence, mortality, survival and incidence-based mortality
 - Regional cancer registries only
 - The huge number of cases, but the representativeness of the data is not the best possible: SEER covers ~25% of US population in 2000–11 (and ~10% in 1973–2011)
 - Measures can be assessed using SEER*Stat Software, see more in <http://seer.cancer.gov/>

Sources of information

- NORDCAN: Incidence, mortality, prevalence and survival statistics in the Nordic countries
<http://www-dep.iarc.fr/NORDCAN/English/frame.asp>
 - Excellent coverage and quality of data
 - Online analyses with graphs, tables and maps for comparisons between
 - 5-year or broader age groups
 - sex
 - countries, and
 - areas within country

Time trends

- **NORDCAN:**

- Time trends available from 1953 onwards, now until 2011 (updated regularly)
- The best source of time series cancer data
- Even if long time series exist and the quality of cancer registration is very high, age-standardized rates may not necessarily be comparable over time and/or between the Nordic countries
 - Look at the patterns of time series data

Predictions

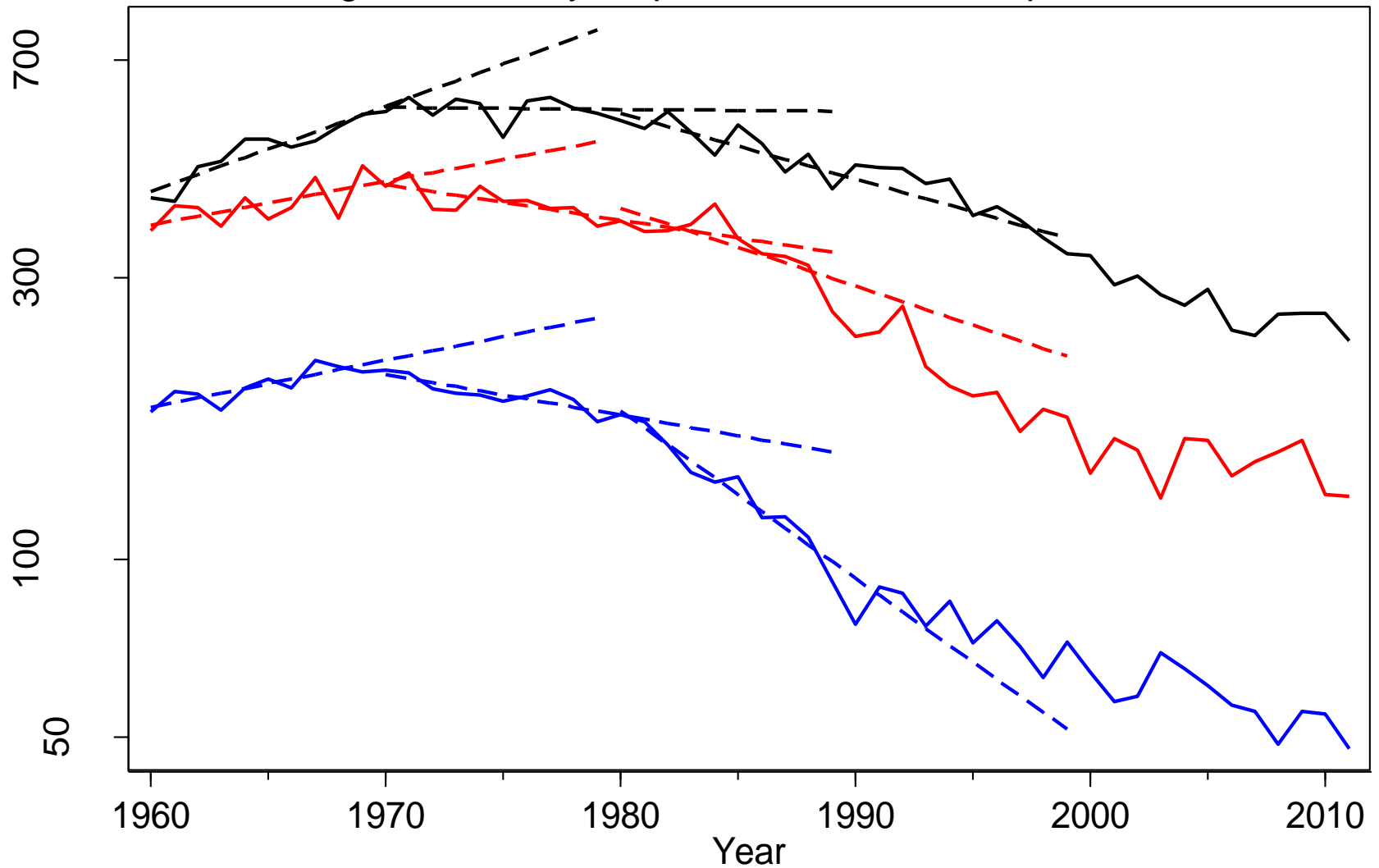
- Supplied by GLOBOCAN and NORDCAN up to 2030s for incidence and mortality
- GLOBOCAN
 - Occurrence in 2012 + future population counts = predicted future numbers
 - Reports future numbers due to incidence/mortality in 2012 and due to demographic factors
 - If the figure for 2012 was estimated, then how much you can rely on an estimate for 2035?

Predictions

- In NORDCAN, short- and long-term predictions are available separately
 - Short-term predictions (up to 5 years)
 - Tend to work if incidence (mortality) has changed (log-) linearly in the prediction base (= a period of the most recent data) prior to prediction

Incidence among 50-59, 60-69 and 70-79 years old men

Lung cancer, 10-year prediction bases and predictions



Predictions

- In NORDCAN, short- and long-term predictions are available separately
 - Short-term predictions (up to 5 years)
 - Tend to work if incidence (mortality) has changed (log-) linearly in the prediction base (= a period of the most recent data) prior to prediction
 - Long-term predictions (after 5 years)
 - Two alternative methodologies, one like in GLOBOCAN and another based on age-period-cohort modelling (NORDPRED)
- Even if the most recent cancer data are reliable, long-term predictions cannot do better than long-term population predictions

Finnish Cancer Registry (FCR)

- Includes diagnoses and deaths from cancer since 1950s
- The coverage of registration is near 100% for malignant and solid tumours
- Publishes cancer statistics
<http://www.cancer.fi/syoparekisteri/en/statistics/cancer-statistics/>
- Non-malignant tumours, such as basal cell carcinoma of the skin, carcinoma in situ in breast and borderline tumour of the ovary, are not included in the statistics
- Publishes statistics on mass screening activities (breast, cervical and colorectal cancer screening programs)

Finnish Cancer Registry (FCR)

- Official statistics presents a big picture, not a detailed information
- If you don't find what you need from our statistics or other publicly available sources, you can send us a specific request for aggregated data by emailing to [kirjaamo\(at\)cancer.fi](mailto:kirjaamo@cancer.fi).
- The better you know, what you want, the easier you get it. Do try to specify your request as well as possible with respect to primary site and histology (ICD-O-3), calendar years, age, sex, and other possible important factors.

Summary

- Cancer is a group of diseases
- Official statistics exist but their coverage and reliability may not be as high as you need
- If you need time series data, NORDCAN is the best choice
- You always need to think behind your data
- Long-term predictions tend to fail

THANK YOU FOR YOUR ATTENTION!