



## **OUTPUTS: PREPARING DATA FOR EXTERNAL USERS, REGISTRY REPORTS.**

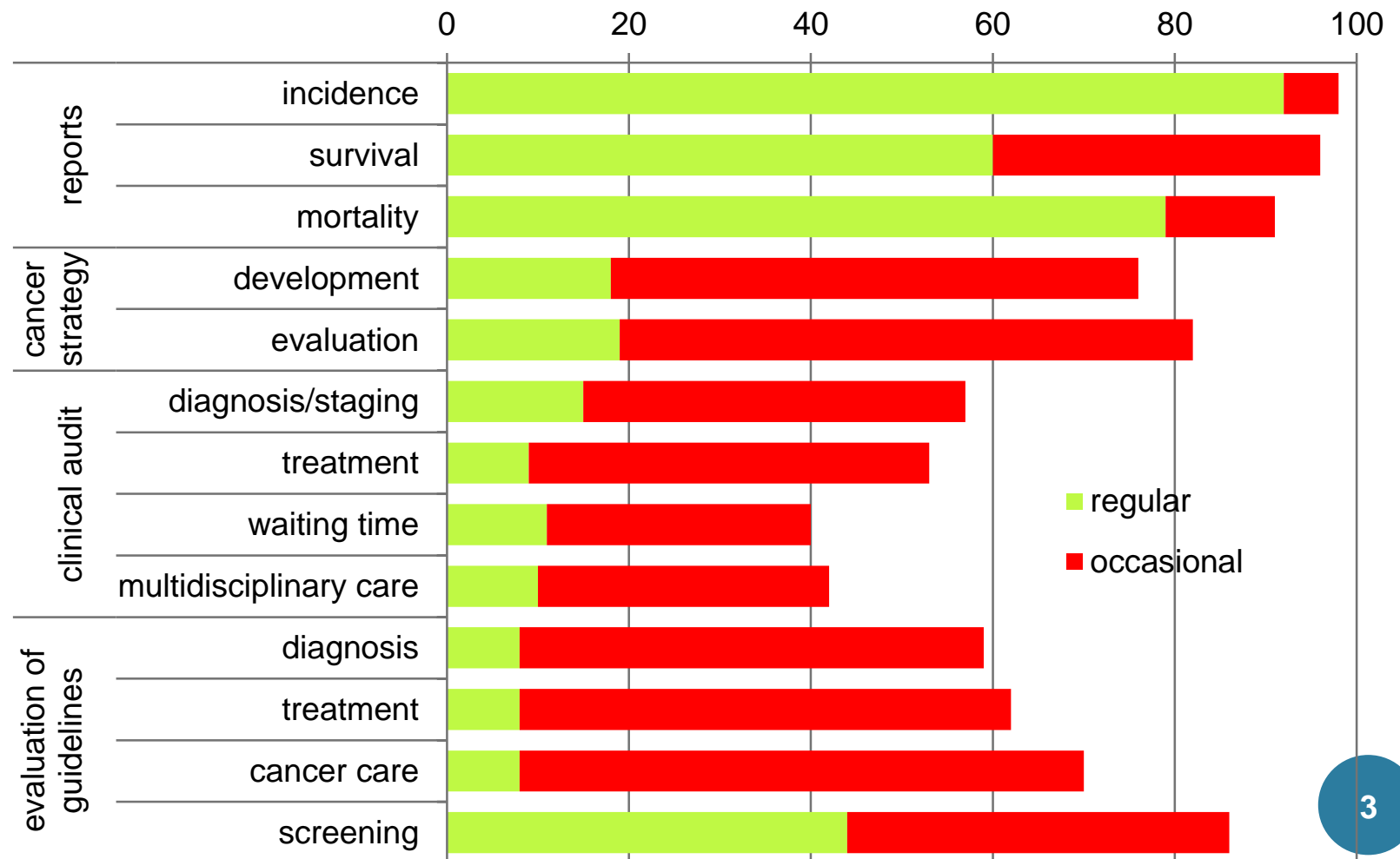
1

**Harry Comber**

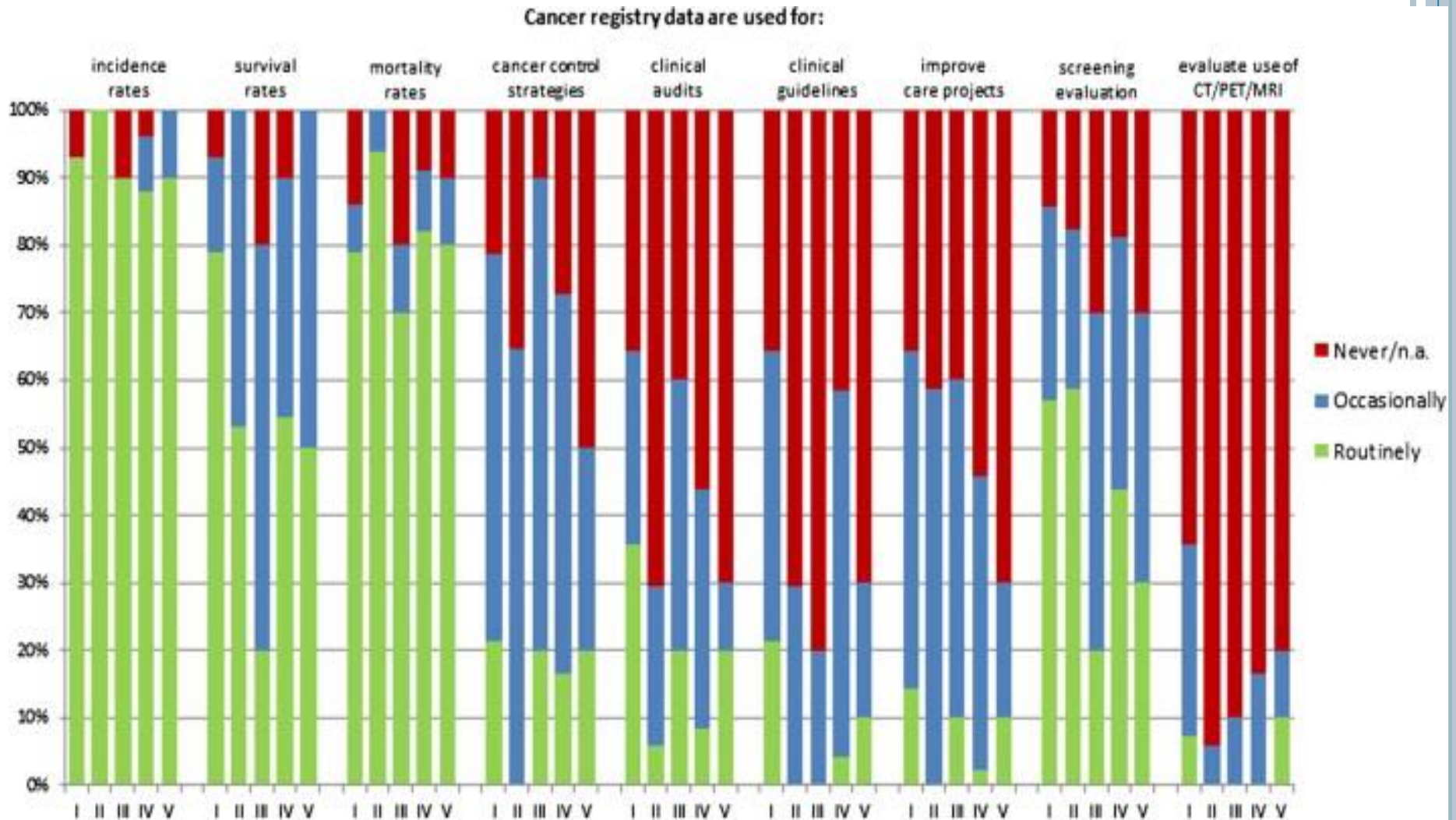
# WHY COLLECT DATA?

- Data not used is worthless
- Dissemination of
  - Publications
  - Data
- What are the aims/duties of the registry?
- Who are the stakeholders?
- Who are the audience?
- What are the messages?

# USES OF REGISTRY DATA



# USES OF REGISTRY DATA



# TYPES OF OUTPUT

- Reports
  - Internal
    - Quality
    - other
  - External
    - Public
    - Other
- Peer-reviewed papers
- Responses to requests
- Data
  - Public
  - Other

# TIMING OF REPORTS

- Annual cycle?
- Timeliness .v. completeness
- Impact

# REGISTRY REPORTS

- Annual report
- Topic-specific
  - Cancer type
  - Childhood
  - Atlas
  - Deprivation etc
- Factsheets

# ROUTINE DATASETS

- Prepared regularly for
  - Website
  - Researchers
  - Other users
- Use of standard software to automate this—SAS, Stata, R



# RESPONSES TO REQUESTS

- Media
- Public representatives
- General public
- Researchers
- “Clusters”

# DATA EXPORT

## International repositories

- IARC
- ENCR
- NORDCAN

## National datasets

- AIRTUM
- GEKID

## Projects

- EUROCARE
- CONCORD

# PREPARING A DATASET

- Identify
  - Variables
  - Years
  - Other inclusion/exclusion criteria
- Extract
- Transform
- Check and clean
- Analyse

# VARIABLES

- What is needed?
- What can be included
  - quality?
  - completeness?

# EXTRACTING DATA

- SQL

- MySQL
- MS SQL Server
- Oracle
- Sybase
- etc

- Data management

- MS Access
- SAS
- Stata
- Etc

# TRANSFORMING DATA

- Data may need to be transformed e.g.
  - Hierarchical data to “flat file” for easier analysis
  - Specific variable order and format for external users
- Dialogue between analysts and data collection on problems

# DATA CHECKING AND CLEANING

- Routine checks e.g. ENCR or IARC checks
- Summaries
- Rare cancers
- Common problems

# ANALYSIS

- Will often throw up inconsistencies
- The best way to identify data problems



# DISSEMINATION

- Thank you

# SESSION 11. STAFF TRAINING

- Does your registry have regular training events for data collection and processing staff?
  - If yes, how often
  - If no, why not?
- How does the registry introduce changes in coding (e.g. in ICD-O, TNM)?
- What training materials do you use
  - Locally generated
  - External—from where?
  - What training materials would you like ENCR to provide?

