Prognostic Outcome Studies

Edwin Chan PhD
Singapore Clinical Research Institute
Singapore Branch, Australasian Cochrane Centre
Duke-NUS Graduate Medical School

Contents

- Objectives
- Incidence studies (Prognosis)
- Risk factor studies (Prognostic factors)
- Clinical prediction rules (Prediction models)

Objectives

- 3 common objectives
  - incidence of outcome studies are follow-up studies describing the incidence risk or rate of an outcome in a defined cohort
  - risk factor studies are interested in measuring the strength of particular prognostic factor(s)
  - risk prediction studies are interested in developing a prognostic model to predict future outcome risk of patients with given levels of risk factors

Incidence studies

- Primary aim is to describe the “burden” of future disease/outcome e.g.
  - the incidence of a disease in an exposed cohort
  - the incidence of an outcome in a treated cohort

- Example:

Objective To determine the short-term risk of stroke and other adverse events after emergency department (ED) diagnosis of TIA.

Specification & Design

- Study Population & Setting
  - demographic, exposure & clinical characteristics
  - representative of target population?
  - homogenous?
- Outcomes to be measured
  - definition
  - follow-up time period
- Design
  - Descriptive cohort study

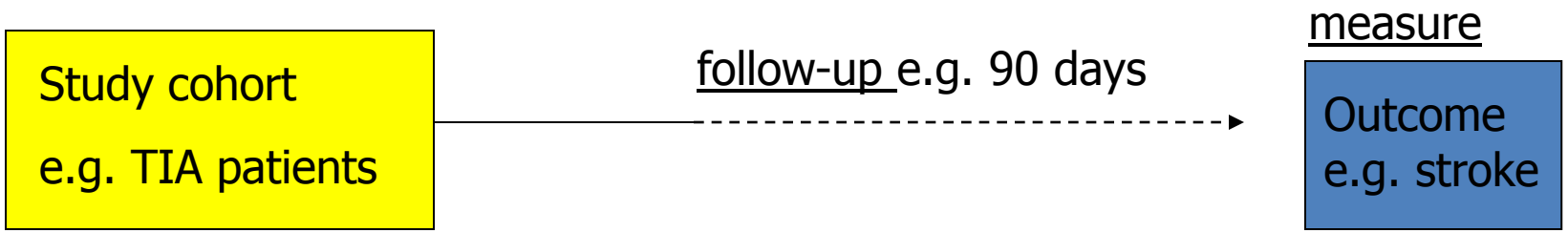
Example

Design and Setting Cohort study conducted from March 1997 through February 1998 in 16 hospitals in a health maintenance organization in northern California.

Patients A total of 1707 patients (mean age, 72 years) identified by ED physicians as having presented with TIA.

Main Outcome Measures Risk of stroke during the 90 days after index TIA; other events, including death, recurrent TIA, and hospitalization for cardiovascular events.

Descriptive cohort study



90-day risk of stroke in a homogenous cohort of elderly patients with TIA

Design issues to take note

- Rare outcomes
  - prospective follow-up is inefficient when the outcome is rare
  - solution: retrospective follow-up
- Unequal follow-up time
  - due to censoring make simple incidence risk estimation invalid
  - solution: use survival analysis

Retrospective cohort studies

Type	Past	Now	Future
Prospective		Study cohort →	Disease occurrence
Retrospective	Study cohort → Disease occurrence	Review past records for cohort & disease details	

Challenges of retrospective studies

- Quality of the data sources
  - missing data
  - inaccurate data
  - data not measured in the relevant way

Validity of risk estimation

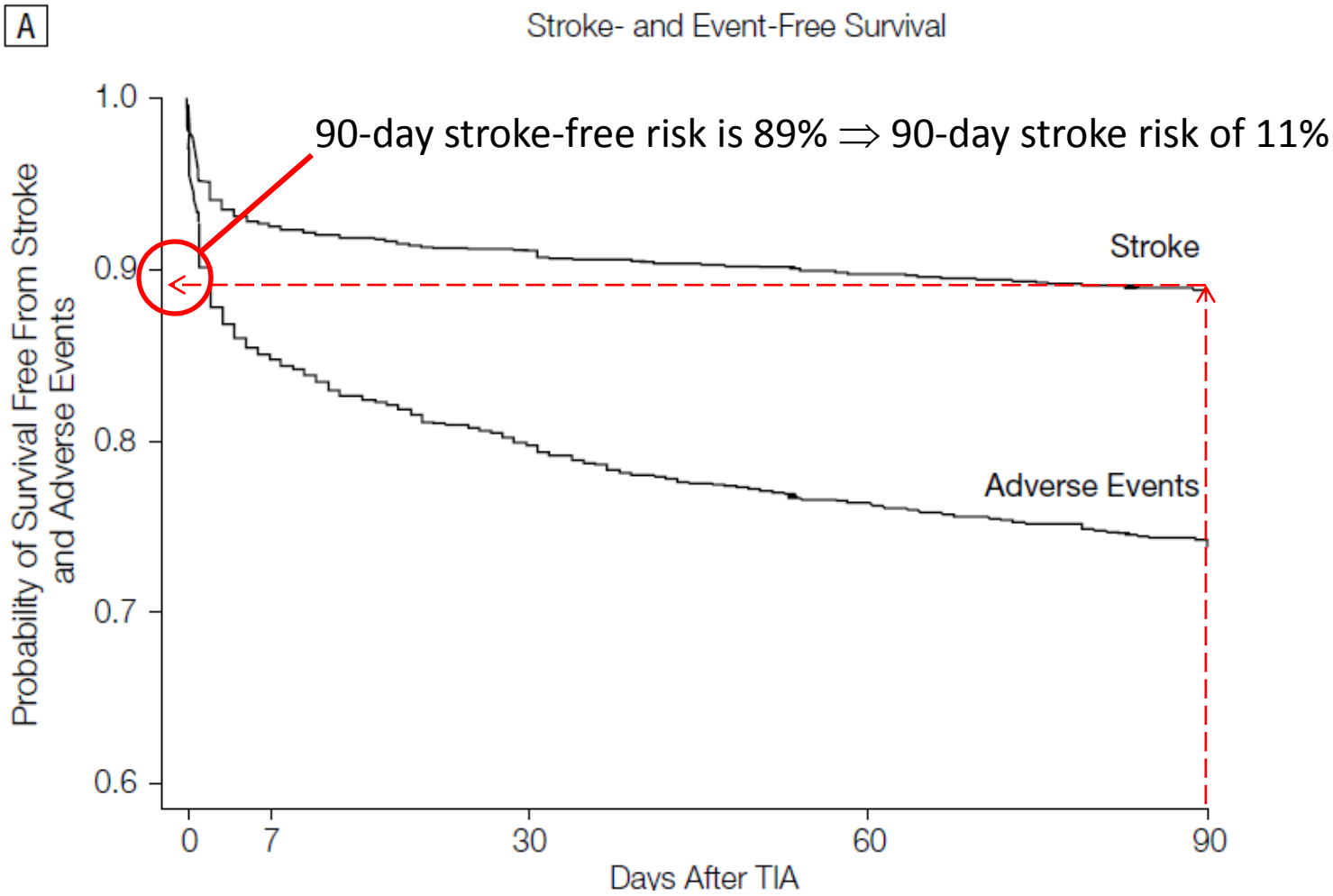
- Direct risk estimation using the fraction of outcomes occurring within the follow-up time is only valid when everyone not having the outcome has the full follow-up time (no censoring)

- Censoring becomes increasingly likely when length of follow-up increases

- An underestimation bias is introduced when it is assumed that all such patients would not have had the event if they had full follow-up

Survival Analysis

Figure. Kaplan-Meier Life-Table Analysis of Survival Free From Stroke and



Design issues to take note

- Cohort definition
  - is the cohort homogeneous for risk?
  - if not, a single risk estimate will not be appropriate
  - solution: investigate other variables that could stratify study cohort into more homogenous subgroups → RISK FACTOR studies
  - Example of risk factors:
    - Gender
    - Age
    - Comorbidities
    - Race
    - Symptom/disease severity
    - Genes
    - Dose of exposure

Study cohort stratified by number of risk factors into 6 subgroups

Table 4. 90-Day Stroke Risk by Number of Risk Factors*

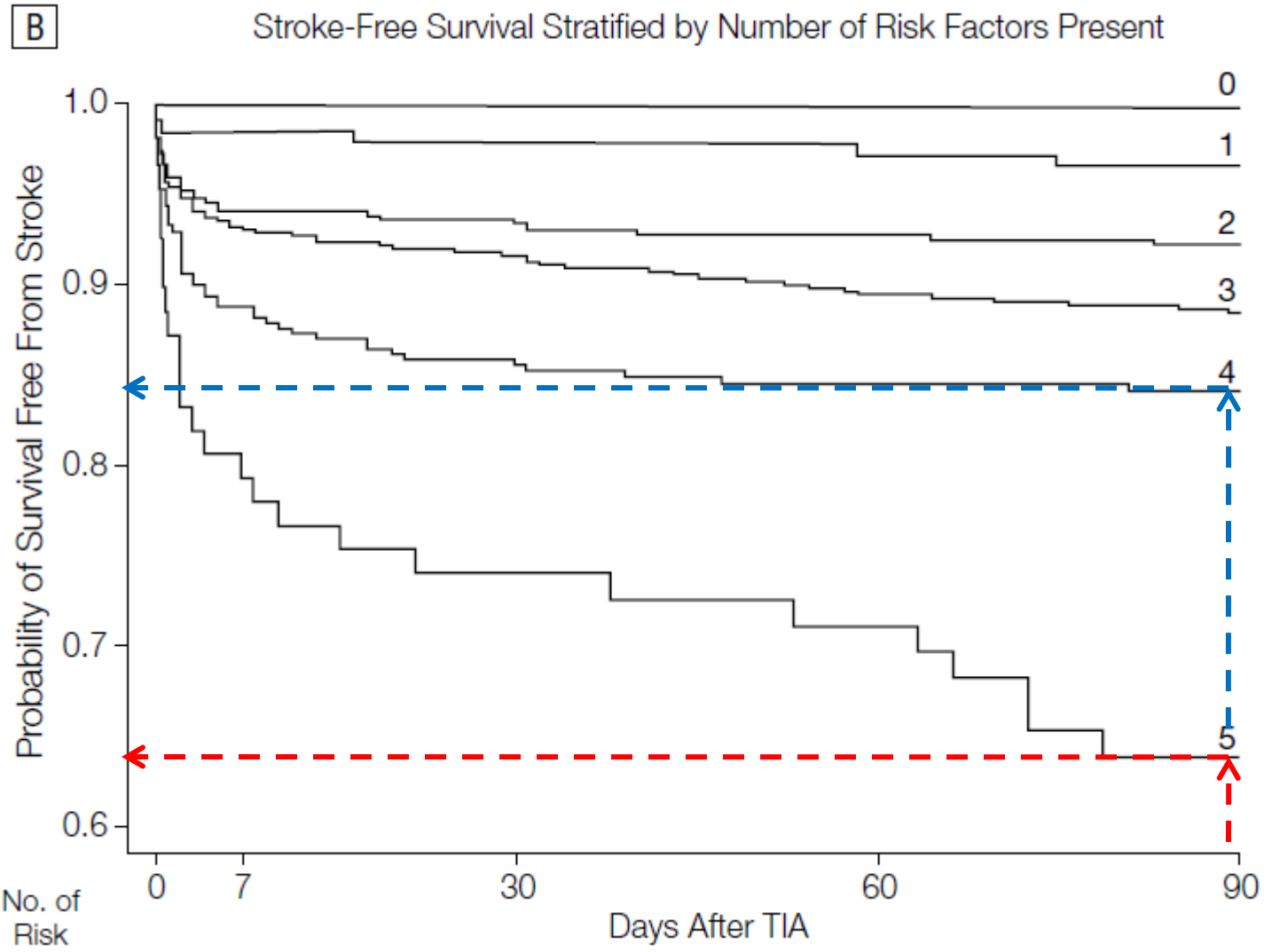
Risk Factors, No.	No. (%) Patients	No. (%) Stroke Within 90 Days
0	22 (1)	0 (0)
1	179 (10)	5 (3)
2	509 (30)	35 (7)
3	584 (34)	63 (11)
4	337 (20)	51 (15)
5	76 (4)	26 (34)

*Risk factors are listed in Table 3.

6 more homogeneous subcohorts

90-day stroke risks

Stratified survival curves



Risk factor studies

- Primary aim is to investigate the comparative effect of exposure versus non- or lower dose exposure to a risk factor

- Study design is a comparative cohort study

- Comparative effect between levels of a risk factor is typically measured as
  - risk, odds or hazard ratio

Table 3. Independent Risk Factors for Stroke Within 90 Days (N = 1707)*

5 risk factors

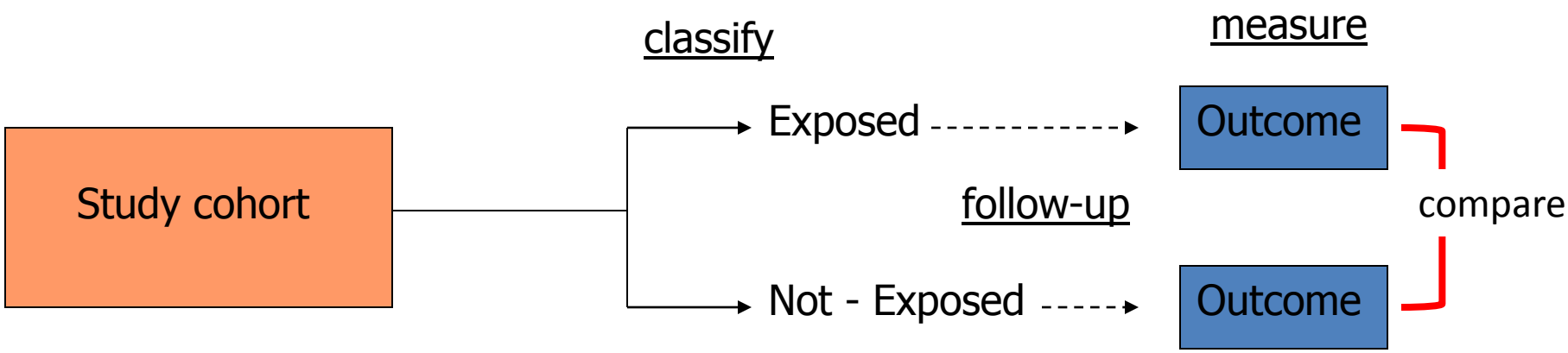
	Odds Ratio (95% CI)	*P* Value
Age >60 y	1.8 (1.1-2.7)	.01
Diabetes mellitus	2.0 (1.4-2.9)	<.001
Duration of episode >10 min	2.3 (1.3-4.2)	.005
Weakness with episode	1.9 (1.4-2.6)	<.001
Speech impairment with episode	1.5 (1.1-2.1)	.01

*Based on logistic regression including all associated variables in univariate analysis ($P<.20$) with stepwise elimination of variables not contributing ($P>.10$). CI indicates confidence interval.

Specification & Design

- Study Population & Setting
  - demographic, exposure & clinical characteristics
  - representative of target population?
- Risk factors/exposures to be measured
  - definition of exposure levels
  - confounders
- Outcomes to be measured
  - definition
  - follow-up time period for which subjects are at-risk because of the exposure
- Designs
  - comparative cohort study
  - case-control study

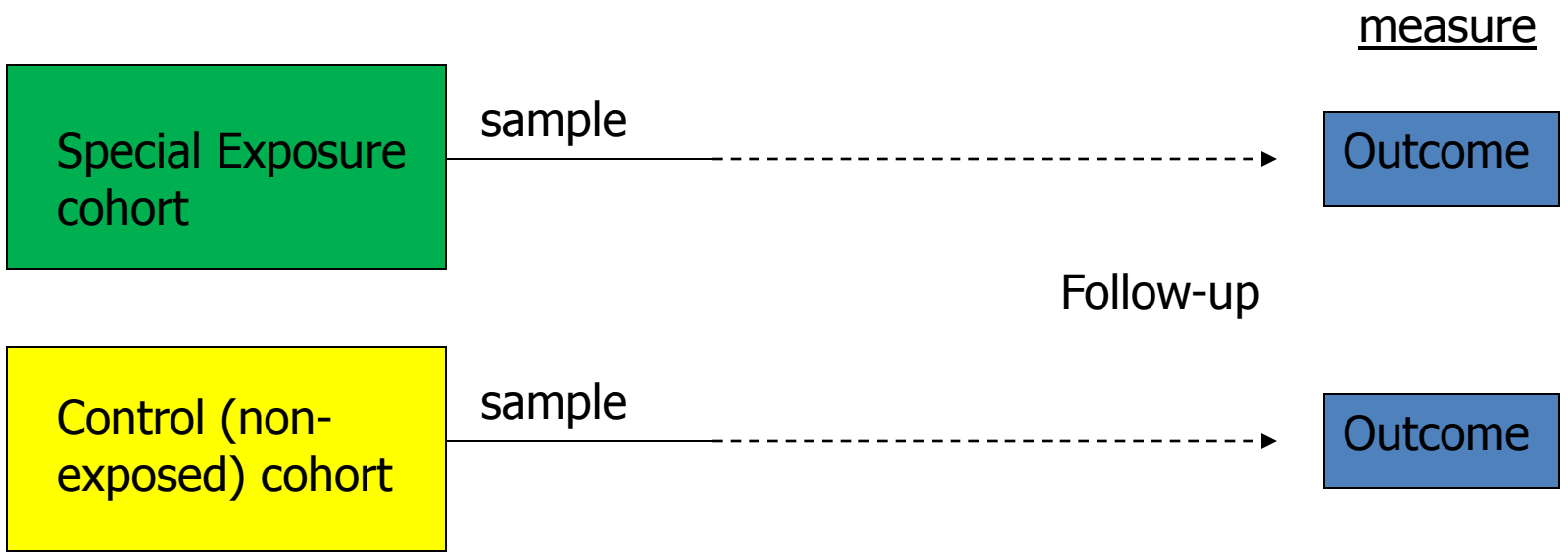
Comparative cohort study



Compare the risk of outcomes between exposed & not-exposed cohorts

Inefficient to conduct such a study when the exposure is rare …

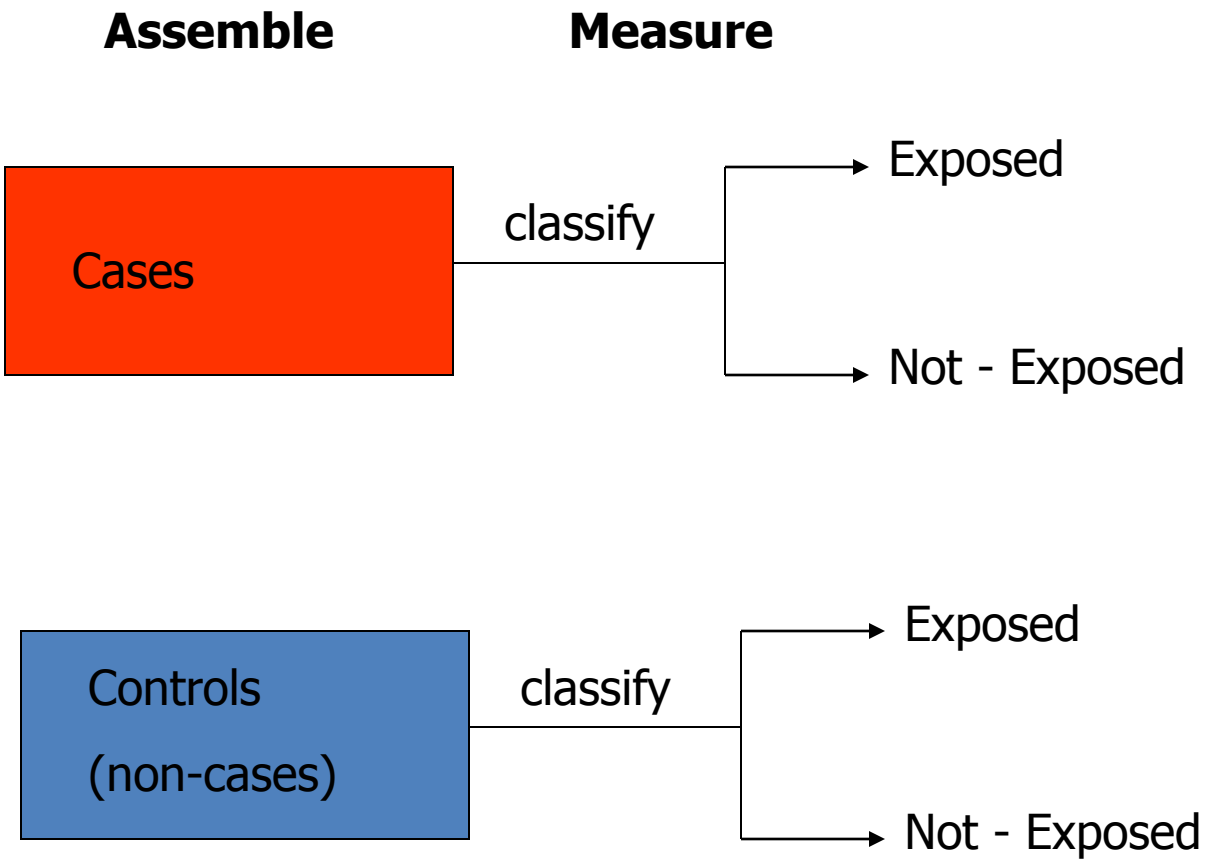
Special exposure cohort study



Design Issues

- Bias in control selection
  - the choice of the control cohort is critical because it must represent the non-exposed in the source population of the special exposure cohort
- Bias in measurement
  - the quality of the measurement for exposure and outcome may be different between exposed and non-exposed cohorts especially in retrospective studies
- Inefficient when outcome is rare …

Case-control studies



Design Issues

- Bias in case selection
  - cases should originate from a single source population & not be a mixture from populations with different causal factors
- Bias in control selection
  - controls should be representative of non-cases in the source population from which the cases were derived from
- Confounder control

Controlling for confounding

- Only for measured confounders
- Stratification – separate analysis within confounder strata
- Standardization – using the same confounder (standard) distribution in each comparator group
- Matching in cohort studies – make the distribution of confounder levels the same in exposed/unexposed
- Adjustment – use multiple regression methods to control for confounders in the analysis, matching in case-control studies can make this more efficient

Stratification

Table 4. 90-Day Stroke Risk by Number of Risk Factors*

Risk Factors, No.	No. (%) Patients	No. (%) Stroke Within 90 Days
0	22 (1)	0 (0)
1	179 (10)	5 (3)
2	509 (30)	35 (7)
3	584 (34)	63 (11)
4	337 (20)	51 (15)
5	76 (4)	26 (34)

*Risk factors are listed in Table 3.

6 more homogeneous subcohorts

90-day stroke risks

Adjustment by multivariable regression

Stroke Within 90 Days (N = 1707)*

	Odds Ratio (95% CI)	*P* Value
Age >60 y	1.8 (1.1-2.7)	.01
Diabetes mellitus	2.0 (1.4-2.9)	<.001
Duration of episode >10 min	2.3 (1.3-4.2)	.005
Weakness with episode	1.9 (1.4-2.6)	<.001
Speech impairment with episode	1.5 (1.1-2.1)	.01

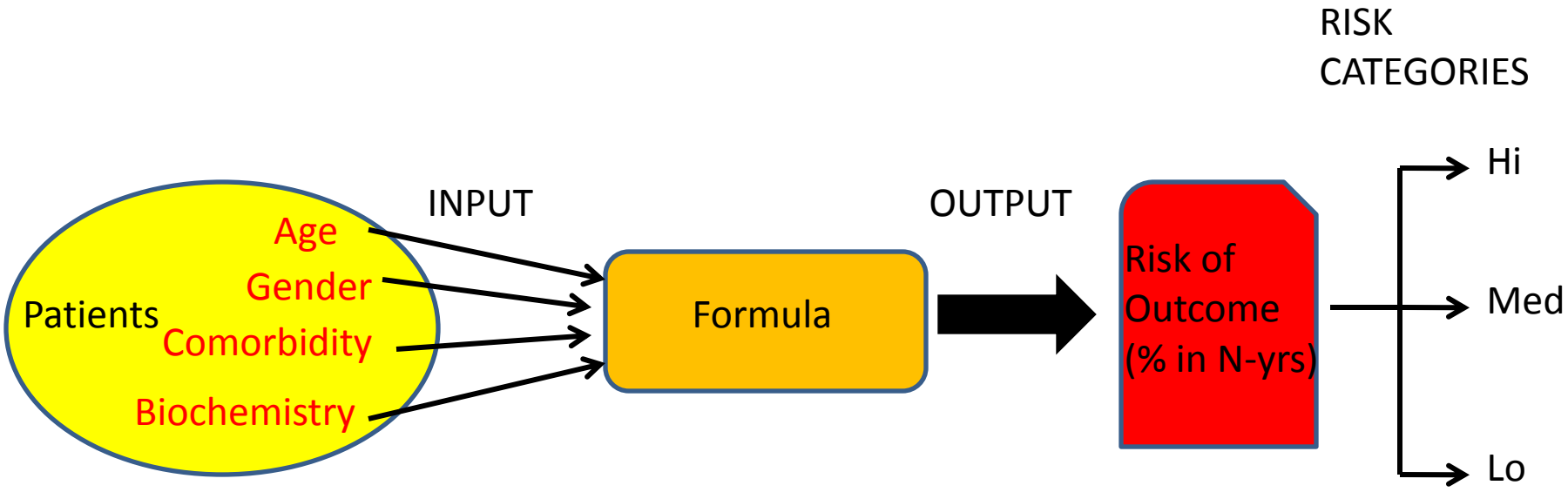
*Based on logistic regression including all associated variables in univariate analysis (*P*<.20) with stepwise elimination of variables not contributing (*P*>.10). CI indicates confidence interval.

Uses of Clinical prediction rules

- Communicate future risk to individual patients
  - more fine-grained risk predictions may be required
- As input into disease management protocols for individual patients
  - may only require classification into broad prognostic risk categories due to limited management options & other over-riding considerations
- To estimate future disease burden for public health planning

Components

- A mathematical formula used to estimate the future risk of an outcome for a specific type of patient from demographic, environmental and disease-specific information (predictors / prognostic risk factors)



Example: Reynolds Risk Score – the Patients

- Aim: To improve on the Framingham Risk calculator for women
- Target Patients: Healthy middle-aged women
- Study Patients: Mainly white women ≥45 y, free of CVD & cancer

Example: Reynolds Risk Score – the Prognostic factors (predictors)

- Candidate factors: 35 demographic, lifestyle, comorbid, lab variables

- Final model factors: 8
  - Age
  - HbA_{1c}%(if diabetic)
  - Current smk
  - Ln(SBP)
  - Ln(HDL-C)
  - Ln(Total cholesterol)
  - Ln(hsCRP)
  - Parental hx of MI <60 y

Checklist for developing valid prediction models I

- General considerations

Step	Issue
Research Question	Strength of predictor or prediction accuracy
Intended application	Clinical/Public Health/Research
Outcome	Clinically relevant
Predictors	Reliable, available, comprehensive
Study design	Retrospective/prospective RCT/cohort/case-control
Statistical model	Appropriate for research question & outcome
Sample size & follow-up time	Sufficient to achieve event numbers

Checklist for developing valid prediction models II

- Modeling steps

Step	Issue
Data inspection	Distribution of values Extent of missing values
Coding of predictors	Predictor transformations & categorization
Model specification	Selection of main effects Testing of assumptions
Model estimation	Shrinkage included?
Model performance	Measures used?
Model validation	Internal validation? External validation?
Model presentation	Appropriate for research question & outcome

Model development

- Study design for data collection
  - cohort vs case-control
    - censoring
  - retrospective vs prospective
    - quality of data

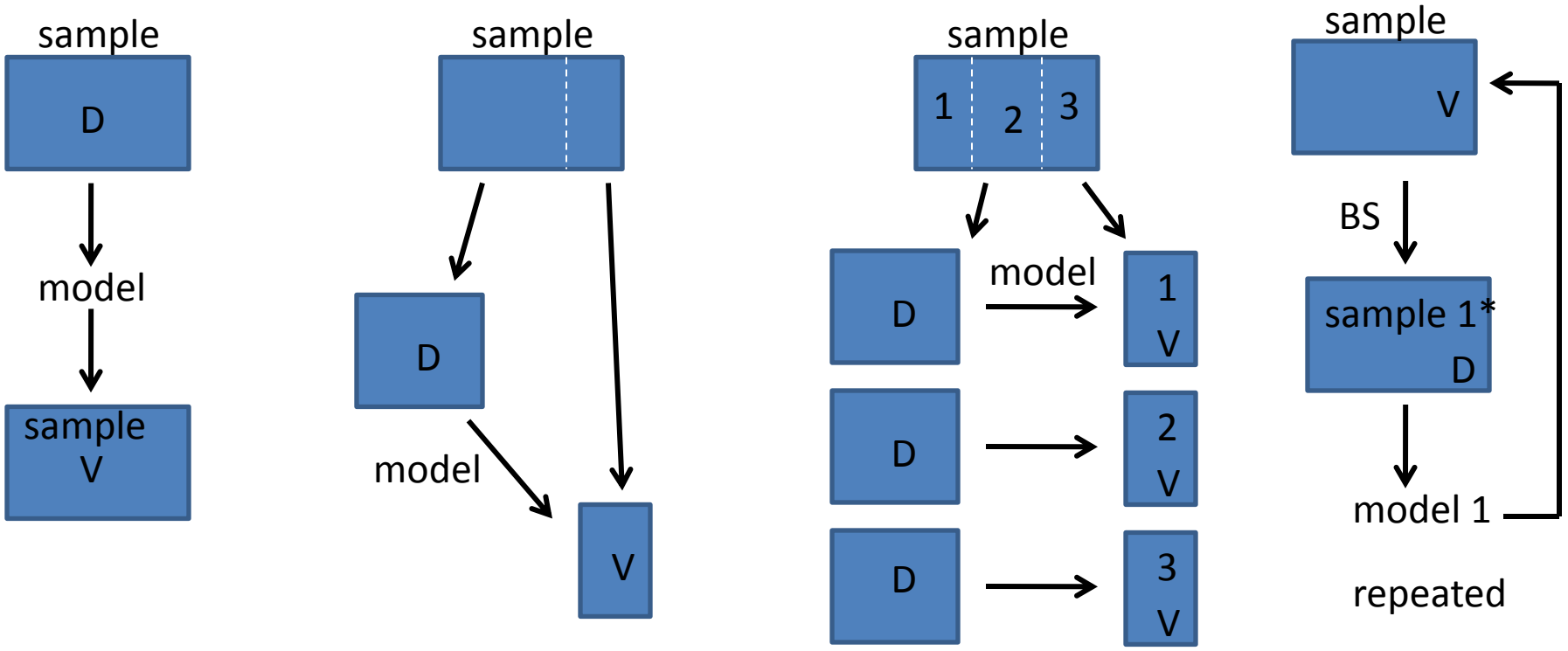
- Data & modeling issues
  - missing predictors
  - distribution of predictor values
  - predictor coding
  - predictor selection (model specification)
  - estimation

Model performance criteria

- Global measures
  - proportion of variability “explained” by the model
  - R^2
- Discrimination
  - how well are cases & controls separated?
  - ROC curve (SE vs SP) & the AUC (*c* statistic)
  - Box plots & the Discrimination slope
- Calibration
  - how close is the predicted to the observed risk?
  - Calibration plot (Observed vs Predicted risk)
    - goodness-of-fit tests

Model validation

- Internal
  - based on development dataset (1 setting)
    - apparent, split-sample, cross-validation, bootstrap



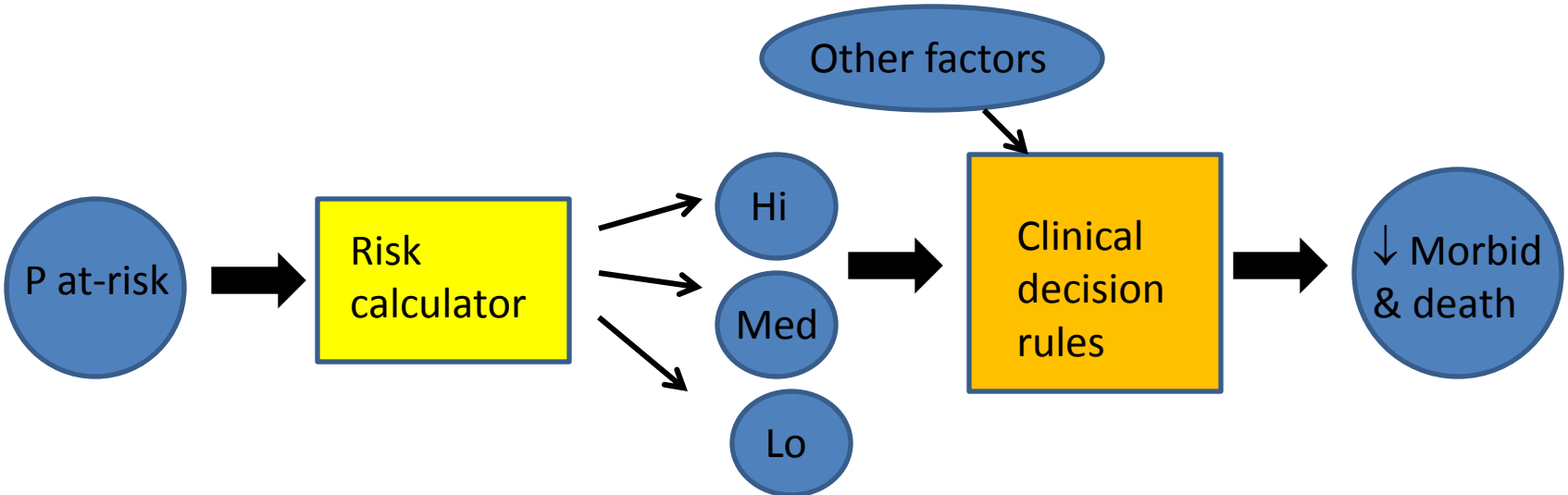
Model validation

- External
  - based on separate plausibly related populations
    - temporal, geographic, fully independent
  - attests to external generalizability (transportability)

- If poor performance …
  - instability of development
  - presence of unaccounted for predictors in ext pop

Impact assessment

- Prediction quality measures are analogous to diagnostic accuracy measures, they do not tell us the impact of actually using the rule compared to not using it. We need to assess the effect of using the risk predictor on clinically important outcomes



Conclusion

- A prognostic model is a mathematical model that relates prognostic factors to outcome prognosis:

Risk (OUTCOME) = f(factor 1, factor 2, …factor k)

Incidence risk

Risk factors

Prediction model

Critical appraisal strategy

- First appraise relevance by considering
  - the question being asked by the authors & the question actually answered by the study reported
  - target population
  - exposure
  - outcome

Critical appraisal strategy

- If relevant, appraise the vulnerability of the study to bias (risk of bias)
  - sampling of study population
  - measurement of exposure
  - measurement of outcome
  - effect of confounders
  - proposed method of analysis

Critical appraisal strategy

- Finally, appraise the study execution, & analysis & reporting
  - the study cohort characteristics
  - quality of the measurements
  - completeness of follow-up
  - control of confounders
  - quantitative reporting