Absolute Risk Prediction

Instructors

Mitchell H Gail, National Cancer Institute

Ruth Pfeiffer, National Cancer Institute

Abstract

Absolute (or “crude”) risk is the probability that an individual who is free of a given disease at an initial age, a, will develop that disease in the subsequent interval (a, t]. Absolute risk is reduced by mortality from competing risks. Models of absolute risk that depend on covariates have been used to design intervention studies, to counsel patients regarding their risks of disease and to inform clinical decisions, such as whether or not to take tamoxifen to prevent breast cancer. This course will define absolute risk and discuss methodological issues relevant to the development and evaluation of risk prediction models. Various study designs and data for model building will be presented, including cohort, nested case-control, and case-control data combined with registry data. Issues relating to the evaluation of risk prediction models and the strengths and limitations of risk prediction models for various applications will be discussed. Standard criteria for model assessment will be presented, as well as loss function-based criteria applied to the use of risk models to screen a population and the use of risk models to decide whether or not to take a preventive intervention that has both beneficial and adverse effects.

Course prerequisites: The course attendees should have a knowledge of basic statistics, epidemiologic designs, and a foundation in survival analysis.

Learning objectives: The attendees of the short course will learn what absolute (or “crude”) risk is, what it can be used for, how to estimate it from data obtained through various designs, and how to assess the usefulness and validity of a model of absolute risk.

Course outline:

1. Definitions and basic concepts
    a. Multiple decrement life table
    b. Hazard function
    c. Competing risk
    d. Pure risk (Kaplan Meier estimate)
    e. Absolute risk (crude risk; cumulative incidence): in survival data competing risk framework
    f. Relative risks

2. Examples of models
    a. Cancer risk: breast, melanoma, colon, lung
    b. Risk of heart disease: Framingham model
    c. Model used after disease diagnosis: prostate cancer
    d. Risk of breast cancer in carriers of mutations in BRCA1 or BRCA2

3. Examples of model applications
    a. Counseling
    b. Designing prevention trials
    c. Clinical decision making for preventive interventions (Tamoxifen)
    d. Assessing burden of risk in population and potential benefits of intervention at population level

4. Model development
    a. Case of no covariates
    b. Flexibility of the cohort data: The Fine-Gray model
    c. Modeling via cause-specific hazards
    d. estimating absolute risk from case-cohort and nested-case control data
    e. Estimating absolute risk from population-based case-control and registry data and associated variance

5. Evaluation of adequacy of model
    a. Standard criteria
        Calibration
        Measures of discriminatory accuracy
        Measures of accuracy
        Prediction error
        Brier statistic with extension to survival data (Schumacher)
    b. Loss-based assessments
        Screening
        Clinical decision-making
    c. Some recently proposed criteria
    d. Comparing two models

6. Additional material in handout:
    Model development: estimating absolute risk from family-based designs
    a. Interpretation of estimands in the presence of special ascertainment and residual familial correlation
    b. Kin-cohort study
    c. Combining relative risks from family-based case-control data with population rates
    d. Variance computations