Meta-Analysis at the Individual Level: *Synthesis of Inferences across Multiple Gerontological Databases*

Karen Bandeen-Roche, Chi Wang
Johns Hopkins Bloomberg School of Public Health

Joint Statistical Meetings
Salt Lake City, UT
July 29, 2007

Acknowledgments

- Johns Hopkins Center on Aging & Health
  - Older Americans Independence Center
  - Women's Health and Aging Studies
  - Linda Fried, Jeremy Walston, Qian-Li Xue
- Colleagues in the present work
  - Dani Fallin, Chris Seplaki
- Luigi Ferrucci, National Institute on Aging

Why synthesize inferences across gerontological databases?

- A remarkable data resource exists
  - Cohort studies on aging: BLSA, Alameda County, EPESE, LSOA, HRS, CHS, SEE, WHAS, Health ABC, PEP, ….
  - Many more: international; focused on cognition; "aging" into relevance
  - Similar measures
- *Statistics: Harnessing the Power of Information!*

Why synthesize inferences across gerontological databases?

- Timely questions in aging need this
  - … to detect subtle risk factor effects
  - … to assess findings' robustness given
    - random sampling variability;
    - systematic sampling distinctions;
    - differential measurement.

Why synthesize: motivation

- Is frailty in older adults caused by systemic biological dysregulation?

- Frailty…
  - A syndrome of decreased "reserve"
  - Many definitions
  - Criteria: Exhaustion, low activity, slowness, weakness, weight loss

- Biological regulation: inflammation
  - A clue to the more complex etiology

Outline

- Sources of variation in inferential targets
- Accounting for (one of) these
- Application: Frailty & inflammation in InCHIANTI; Women's Health & Aging Study (WHAS)
- What's new?
  - Synthesis across challenges
  - Conceptual framework needed to accomplish this

Sources of variation

- Variation about what target?
  - Superpopulation: S
  - A population characteristic: Φ
  - If Bayesian: prior belief $F(\Phi)$—source 1

- Designs $D_h^{(n)}$ are employed to sample S
  - *h* indexes study; *n* indexes sample size
  - Study *h* targets parameter $\Phi_h \sim F(\Phi_h|\Phi)$—source 2
    - "Random" variation: incidental conduct differences
      - Most prior work here
    - "Systematic" variation: selection differences

Sources of variation

- Variation about what target?
  - Superpopulation: S
  - A population characteristic: Φ
  - Sources 1 & 2: $F(\Phi)$; $F(\Phi_h|\Phi)$

- There is variability sampling within study *h*
  - Individuals $i=i_1,\ldots,i_n$ chosen in a particular application of design $D_h^{(n)}$
  - Let "μ_{hi}" represent perfectly measured data
  - Source 3: $F(\mu_{hi}|\Phi_h,\mu_h)$; $T(\mu_{hi}) = \hat{\Phi}_h \xrightarrow{p} \Phi_h$

Sources of variation

- Variation about what target?
  - Superpopulation: S
  - A population characteristic: Φ
  - Sources 1-3: $F(\Phi)$; $F(\Phi_h|\Phi)$; $F(\mu_{hi}|\Phi_h,\mu_h)$

- Data may be imperfectly measured
  - More importantly: differentially across studies
  - Data "M_{hik}" measure μ_{hi}, $k=1,\ldots,K_h$
  - Source 4: $F_h(\underline{M}_{hi}|\mu_{hi})$

Inoue et al., Biostatistics, 2004

Accounting for multi-source variation

- Option 1: big hierarchical model
  - Common; well used
  - Difficulties
    - Specifying the component distributions
    - Few studies

- Option 2: tackle per source, then synthesize

Accounting for multi-source variation
Top-down

- Measurement error: possibly really hard
  - In application:
    - Three criteria quite differently measured
    - Similar criteria were differently, oppositely prevalent
    - Outcome for this talk = walking speed
  - Unless otherwise noted: $M_{hi}=\mu_{hi}$
- Sampling variation: Usual means
- Inter-study variation
  - Population composition
  - Random

Accounting for multi-source variation
Differential population composition

- Key reference: ***Elliott & Davis, Appl Stat, 2005***

- Notation
  - *i* identifies a person in the superpopulation
  - Implement / choose a subset of designs-here, 2
  - S_i codes the study, *h*, into which *i* is sampled
    - Say, *a*=InCHIANTI; *b*=WHAS; *c*=neither
  - We aim to estimate $F(Y_i|\underline{X}_i,\Phi,\mu)$ in S ; $(Y,\underline{X})=M$

- Study *a* identifies $F(Y_i|\underline{X}_i,S_i=a)$

$$= Pr(S_i=a|Y_i,\underline{X}_i)F_S(Y_i|\underline{X}_i)/Pr(S_i=a|\underline{X}_i)$$

Problem: We typically don't know $Pr(S_i=a|Y_i,\underline{X}_i)$, $Pr(S_i=a|\underline{X}_i)$

Accounting for multi-source variation
Differential population composition

- One alternative to

$$F_S(Y_i|\underline{X}_i) = F(Y_i|\underline{X}_i, S_i=h)Pr(S_i=h|\underline{X}_i)/Pr(S_i=h|Y_i,\underline{X}_i) \quad (1)$$

> Compare to a "reference" study:

If (1) is satisfied for h=a,b, then $F(Y_i|\underline{X}_i, S_i=a)$ =

$$[odds(S_i=a:b|Y_i,\underline{X}_i)/odds(S_i=a:b|\underline{X}_i)]F(Y_i|\underline{X}_i, S_i=b) \quad (2)$$

- Estimate leading factor (say, logistic regressions)
- Use as weights (i) per or (ii) pooling studies
- If (i): Combine per-study estimates as last step

Accounting for multi-source variation
Differential population composition

- What does it mean for (1) to be satisfied for h=a,b?
  - Application: Association between inflammation and frailty same in WHAS, InCHIANTI target populations
    - Controlling for measured covariates
    - i.e. the only "issue" is population mix re Y and X
    - Concern: very different cultures
  - What does it mean to "sample the same population", beyond population mix?
  - Definition 1: There exists D
    - with identical support in both studies
    - such that $F(y|D_i,x,S_i=a) = F(y|D_i,x,S_i=b)$ a.e. (y,x).

Accounting for multi-source variation
Differential population composition

- Now, $F(Y_i|\underline{X_i}, S_i=a)$ may be expressed as

$$E\{[odds(S_i=a:b|D_i,\underline{X_i})/odds(S_i=a:b|\underline{X_i})]F(Y_i|D_i,\underline{X_i},S_i=b)\}$$

- If D_i partially accounts for non-representative sampling of Y, we recommend

$$E\{[odds(S_i=a:b|Y_i,D_i,\underline{X_i})/odds(S_i=a:b|\underline{X_i})]F(Y_i|D_i,\underline{X_i},S_i=b)\}$$

- In both cases, "E" is with respect to $F(D_i|X_i,S_i=b)$, but weighting conforms it to $F(D_i|X_i,S_i=a)$

Application
Data (InCHIANTI: n=200; WHAS: n=682)

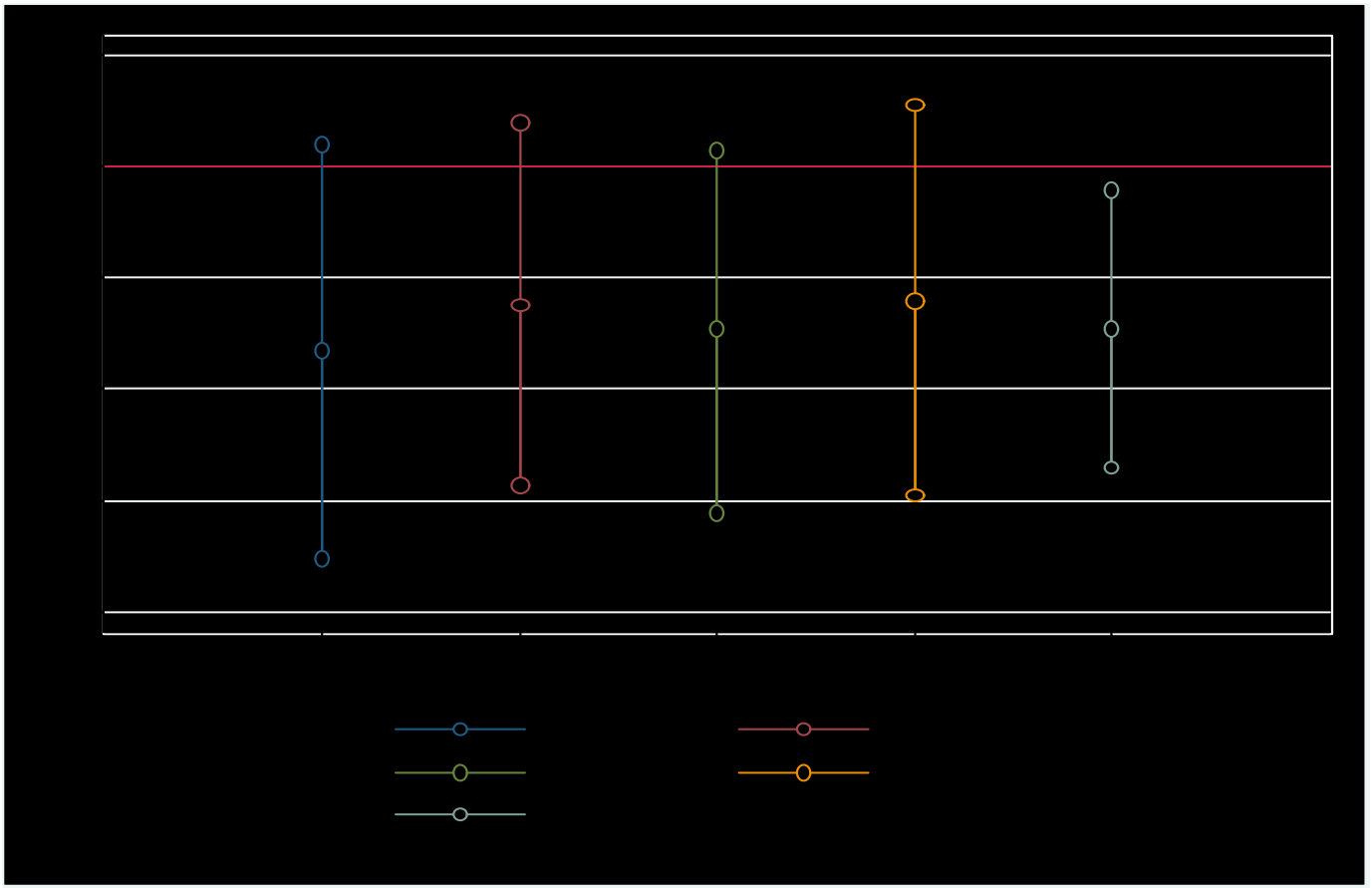
- Y = walking speed
  - Means = 0.98, 0.89; SDs = 0.19, 0.33
- Adjustment covariates
  - Ever smoked: 18.0%, 47.7%
  - Inflammatory disease: 22.5%, 44.2%
  - Age: 70-79 years; women only
- D = prevalent mod/severe disability
  - Partitions WHAS into two **separate** studies
  - Disabled: 27.0%, 40.9%
- Primary covariate: log IL-6 conc.
  - Meas. 1: Geom. Means = 1.34, 2.83
  - WHAS Meas. 2: Geom. Mean = 3.35

Application
Analysis: Speed and inflammation

- InCHIANTI = reference study
- Adjustment of WHAS to InCHIANTI
  - = ratio of odds (per person)
  - estimate each by logistic regression
  - outcome = "study" (**1**{InCHIANTI})
  - numerator: predictors = D, x, (Y)
  - denominator: predictors = x
  - Range of estimated weights:
    - Disability-adjusted: 0.83 to 1.33
    - Fully-adjusted: 0.59 to 2.84

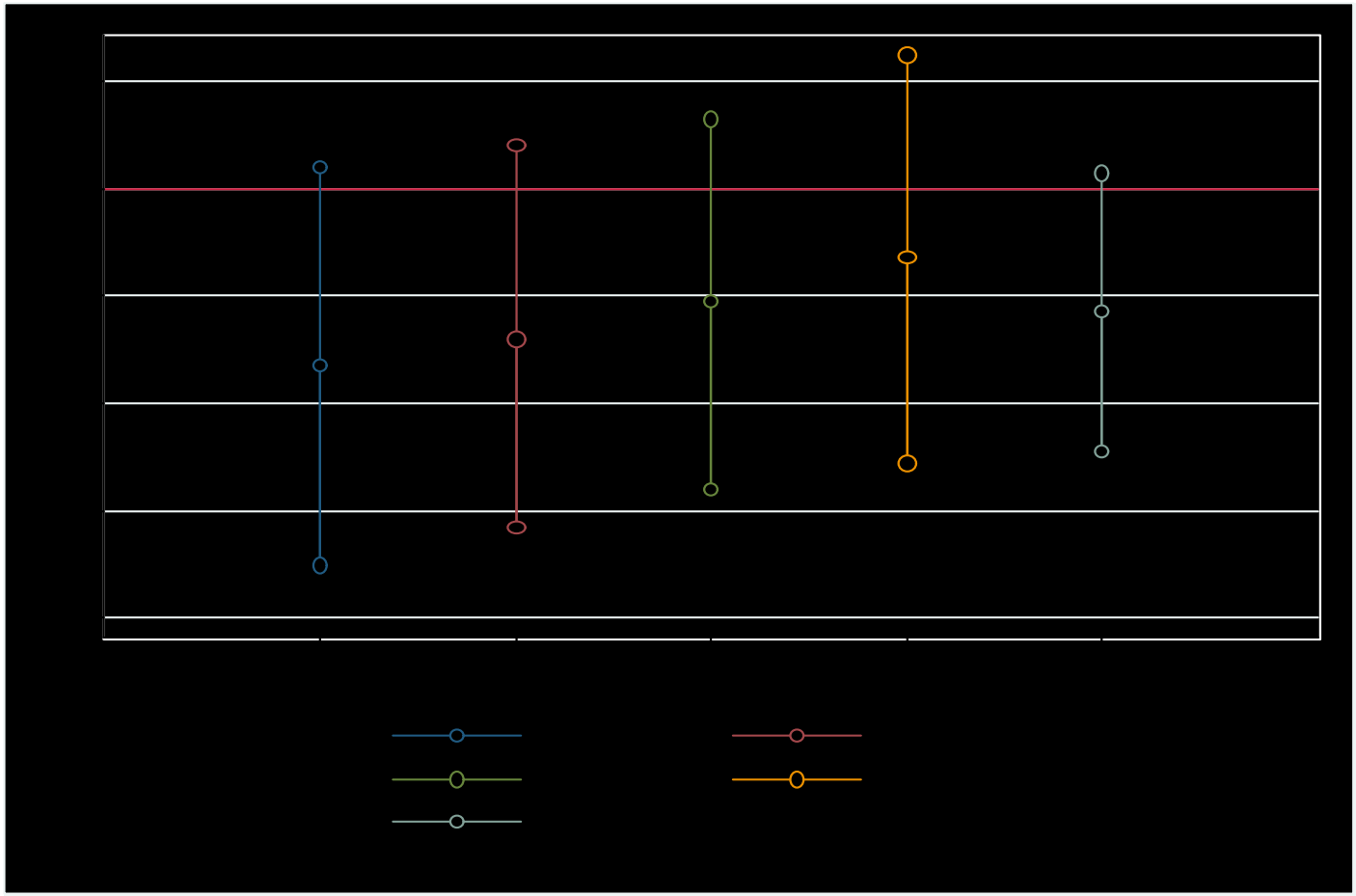
Application
Analysis: Speed and IL-6 (corrected)

95% CIs

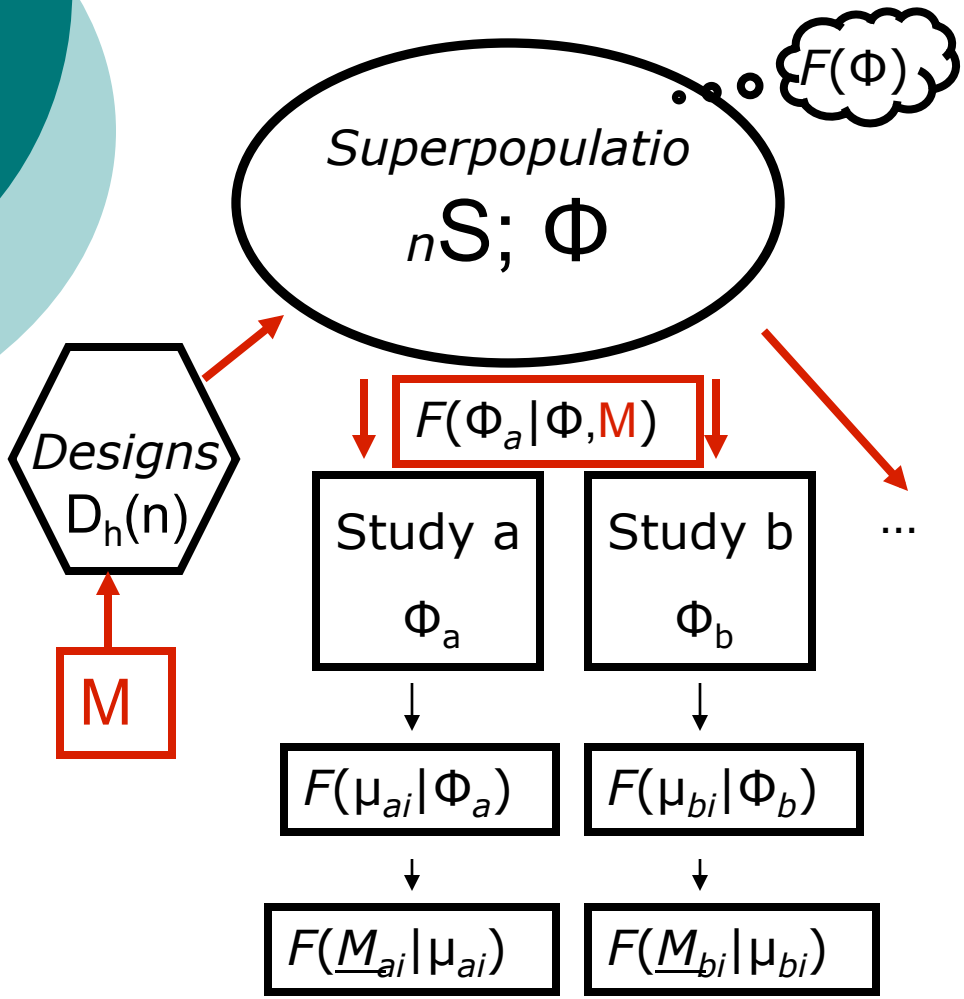


Application
Analysis: Speed and IL-6 (original)

95% CIs

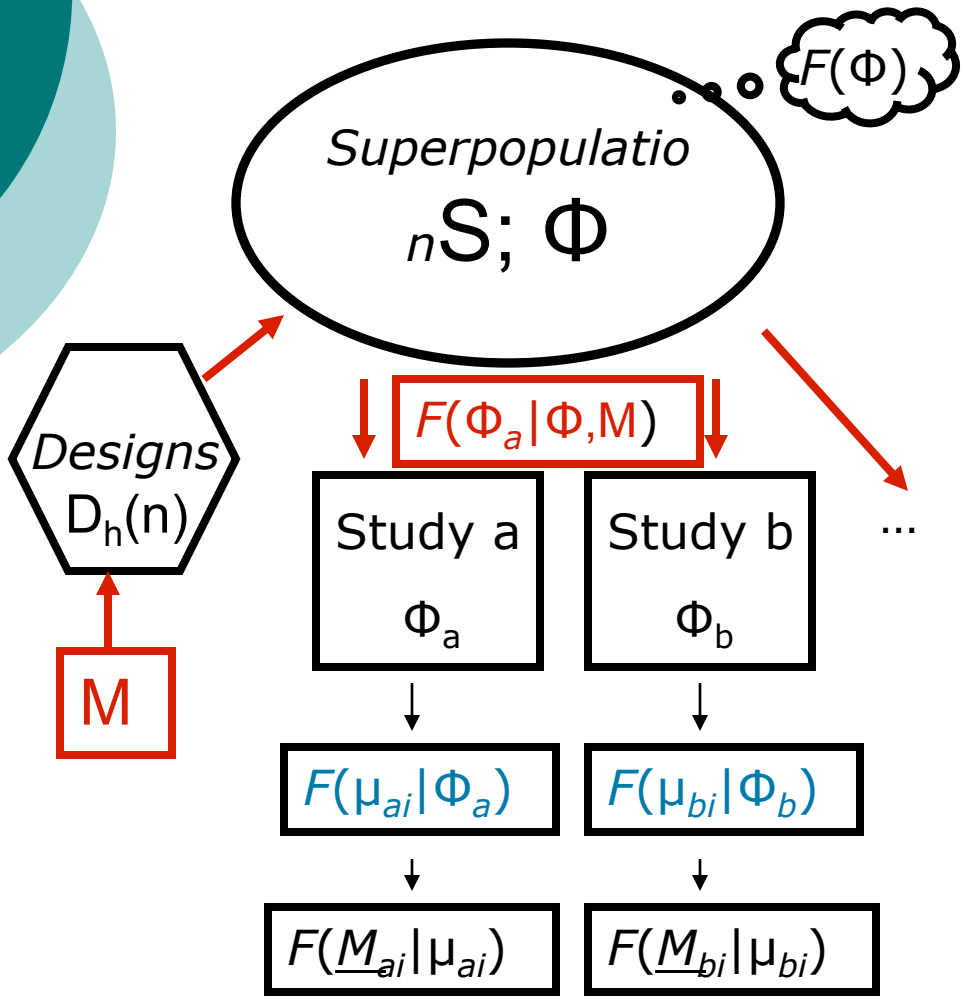


Discussion
Inferential accounting for variation



- For a few studies chosen for availability: fixed effects?
  - Data application: no substantial between-study heterogeneity

Discussion
Inferential accounting for variation



- From a random study-sampling point of view: need
  $E[(\hat{\Phi}_h-\Phi)^2|M]$
  = within study variance σ^2_h + between study variance, σ^2_M
- With two studies:

$$\hat{\sigma}^2_M=(\hat{\Phi}_a-\hat{\Phi}_b)^2+\hat{\sigma}^2_a+\hat{\sigma}^2_h$$

Discussion

- A first step
- Issues needing deeper solutions
  - Integration of hierarchical, weighting approaches
  - Flexible modeling in weighting approach
  - Accounting for variability in estimation of weights
  - Collapsibility
  - Mutual referencing, rather than to one study
  - Delineation of extent to which superpopulation inferences can be made; implications for design
- Implication: groundwork toward more valid synthesis of findings from multiple epidemiological studies