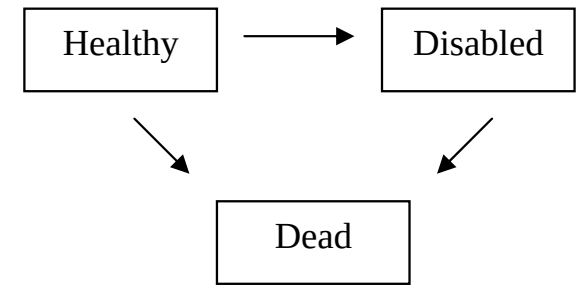
Two Tools for the Analysis of Longitudinal Data: Motivations, Applications and Issues

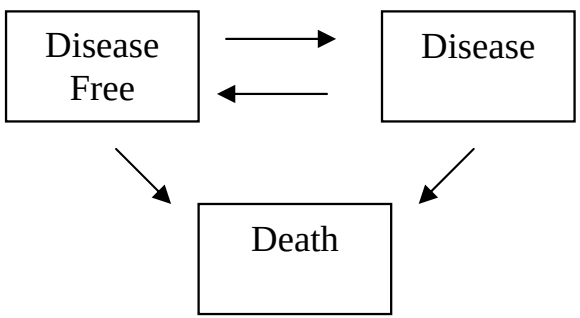
Vern Farewell

Medical Research Council Biostatistics Unit, UK

Flexible Models for Longitudinal and Survival Data
Warwick, UK
July 27, 2015

Introduction





- Subjects/patients observed repeatedly over a period of time.
- When seen, can be classified into different *states*.

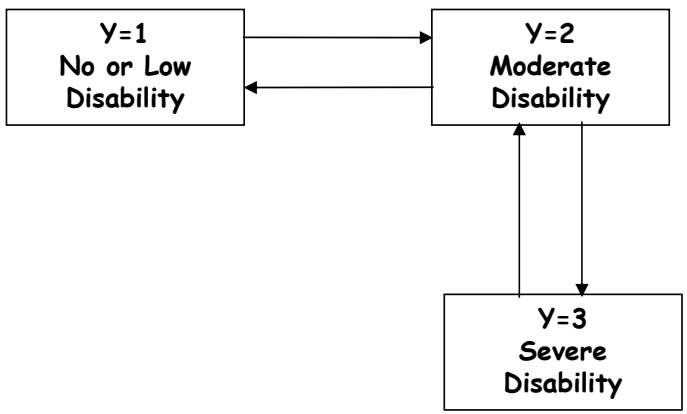
Estimation

- Transition rates typically modelled as relative risk regression models.
  - $\lambda_{(a,b)i}(t \mid x_i(t)) = \lambda_{(a,b)0}(t) \exp(\beta' x_i(t))$
  - Transition rate from state a to b for subject i
- Maximum Likelihood Estimation
  - Continuous time: Individual contribution is the probability of an observed transition between two time points which is a sum of the possible pathways if there are unobservable states.
  - Intermittent Observation: Contribution is probability of the observed state change between two time points of observation. May involve standard calculation of transition probabilities for stochastic processes or more specialised computer code (R).

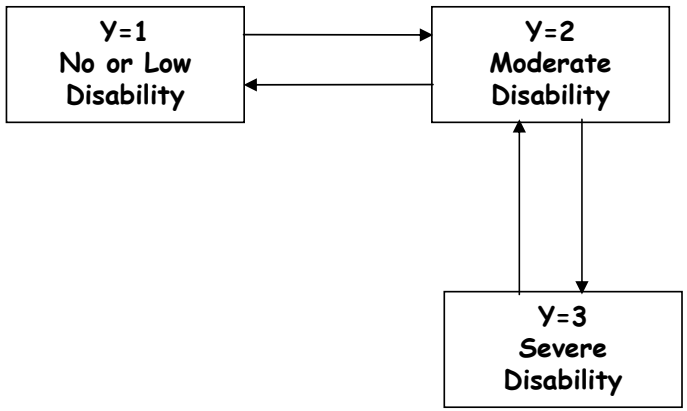
Psoriatic Arthritis (PsA)

- Inflammatory arthritis associated with psoriasis
- Joint involvement is similar to rheumatoid arthritis but general patterns differ
- Disease activity is reflected in joints being swollen (*effused*) and/or painful (*tender*)
  - Activity is reversible
- Disease progression is taken to be reflected in damaged joints
  - Damage is generally held to be irreversible
- Quality of Life - functional disability [HAQ]
- Six monthly data on 600 patients entering a clinic since 1978 (HAQ collected annually since 1993)

Three-state Model for HAQ States



Findings from three-state model for HAQ



- Patients spend twice as long, on average, in the no disability state than the others.
- 46% did not change states.
- 27% changed in only one direction.
- 27% both improved and worsened.

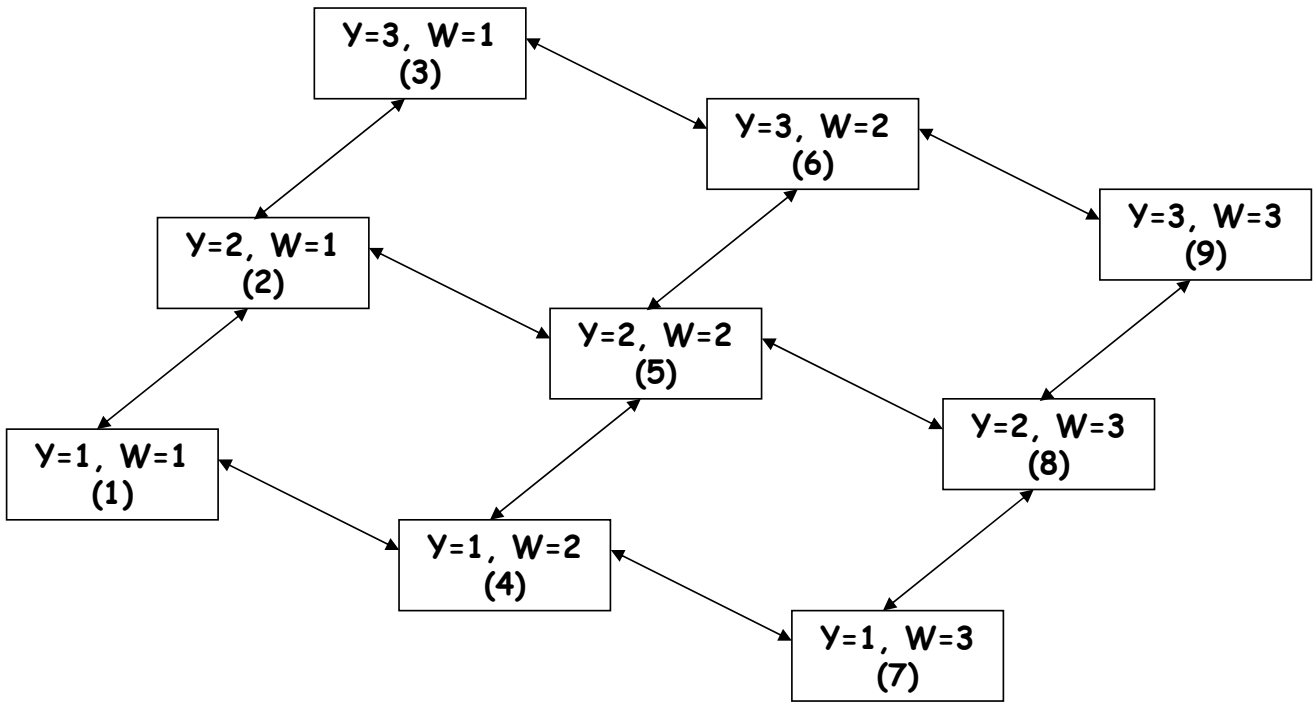
Regression modelling

- Important to adjust for disease activity which is rapidly fluctuating time-dependent ordinal variable.
  - Three States: 0, (1-5), 5+ joints swollen or painful [$W(t)$: Will generate 2 binary indicator variables in three-state model.]
- two other variables of direct interest:
  - Number of damaged joints (less changeable time-dependent variable) [$X(t)$]
  - Sex (time-independent) [Z]
- Use relative risk regression modelling of transition rates between states.
- Only observe patients intermittently at the times of patient visits.

Assumptions

- To fit this model we need to make assumptions
- Panel data generates intermittent observation of states and explanatory variables
- Typically assume transition rates are piecewise constant:
  - Baseline rates piecewise constant
  - Time-dependent explanatory variables assumed constant between clinic visits
- To assess second assumption, we jointly modelled activity and disability
  - Nine-state model created for activity and disability
  - Constraints on baseline intensities and regression parameters required to ensure equivalence with three-state disability model of interest

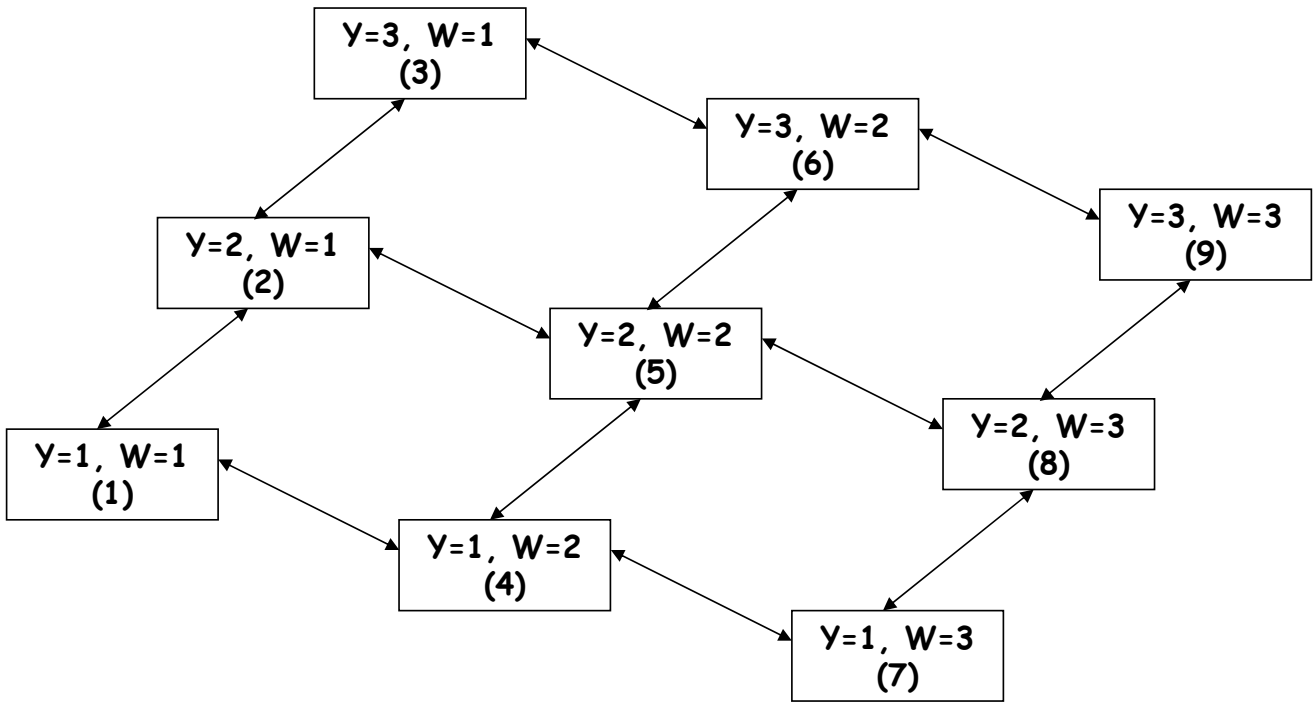
Nine-state Model



Time-dependent explanatory variables assumed constant between clinic visits

- Essentially a situation of measurement error
- Large literature on measurement error in explanatory variables
  - Errors often have mean zero
  - Attenuation is usually seen
  - Some work in survival analysis on infrequently updated time-dependent explanatory variables
- Here interest in mismeasured variable is for adjustment purposes
- Assume interest is in relationships not prediction when use of observed variables may be more sensible

Nine-state Model



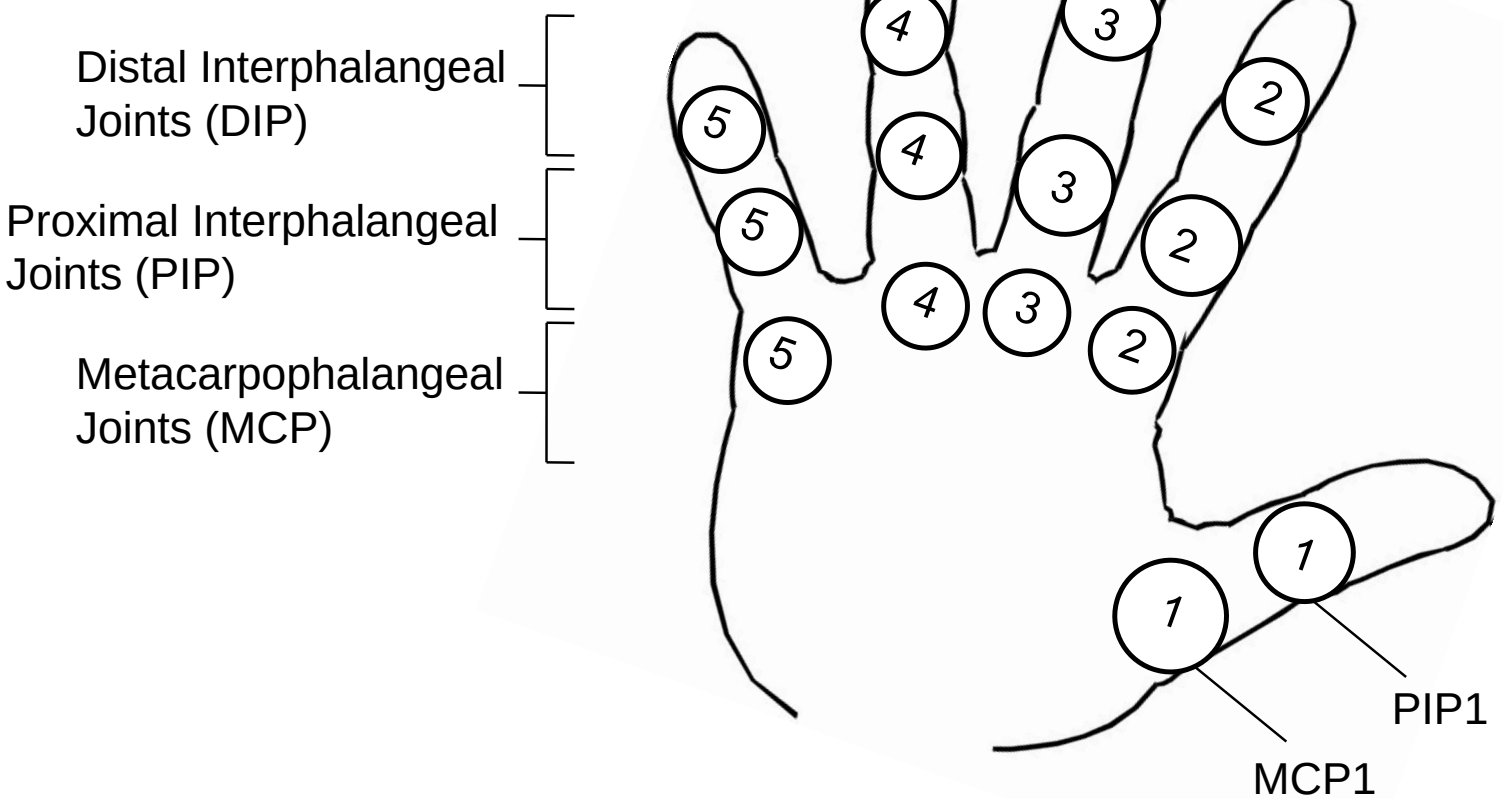
- Constrain the regression coefficients in the rates for all upward transitions between the same two Y (HAQ) states to be same and, similarly, for downward transitions.
- Transition rates, specified by the baseline transitions rates and the regression coefficients, between the same two W (Activity) states are constrained to be identical.
  - Need to model the "marginal" distribution of $(W(t-)|X(t-), Z)$ to get at the joint distribution, $(Y(t)|W(t-), X(t-), Z)$.
- The baseline transition rates between states of the $Y(t)$ sub-process are unconstrained to model the dependence of $Y(t)$ process on $W(t-)$.
- Explanatory variables permitted to modify baseline transition rates for the $W(t)$ sub-process. Handles confounding and mimics the usual regression formulation where no assumption of independence between explanatory variables is made.

Variable	Disability Transition	Multi-state Representations: Misspecified Model Estimate (95% CI)	Multi-state Representations: Expanded Model Estimate (95% CI)
Sex			
Male v Female	$1 \to 2$	$-0.7578(-1.0350, -0.4801)$	$-0.6444(-0.9315, -0.3573)$
Male v Female	$2 \to 3$	$-0.1977(-0.6329, 0.2376)$	$-0.2483(-0.6911, 0.1945)$
Male v Female	$2 \to 1$	$0.0931(-0.1749, 0.3610)$	$0.1640(-0.1110, 0.4391)$
Male v Female	$3 \to 2$	$0.1506(-0.2506, 0.5518)$	$0.07219(-0.3393, 0.4837)$
Number of Damaged Joints			
	$1 \to 2$	$0.0106(-0.0019, 0.0231)$	$0.0103(-0.0025, 0.0230)$
	$2 \to 3$	$0.0036(-0.0112, 0.0183)$	$0.0086(-0.0070, 0.0242)$
	$2 \to 1$	$-0.0166(-0.0273, -0.0058)$	$-0.0201(-0.0307, -0.0095)$
	$3 \to 2$	$-0.0116(-0.0244, 0.0013)$	$-0.0168(-0.0302, -0.0033)$
Number of Active Joints			
$[1, 5]$ v 0	$1 \to 2$	$0.4892(0.1774, 0.8009)$	$1.0283(0.4972, 1.5594)$
$[1, 5]$ v 0	$2 \to 3$	$0.2943(-0.3530, 0.9416)$	$1.1029(-0.2906, 2.3491)$
$[1, 5]$ v 0	$2 \to 1$	$0.1004(-0.2576, 0.4584)$	$0.1429(-0.3331, 0.6188)$
$[1, 5]$ v 0	$3 \to 2$	$-0.1865(-0.8610, 0.4879)$	$-0.2847(-1.2181, 0.6486)$
> 5 v 0	$1 \to 2$	$0.7924(0.4269, 1.1580)$	$1.6063(1.1409, 2.0716)$
> 5 v 0	$2 \to 3$	$0.7286(0.1158, 1.3410)$	$1.7995(0.6943, 2.9047)$
> 5 v 0	$2 \to 1$	$-0.0045(-0.3605, 0.3515)$	$-0.9484(-1.4376, -0.4592)$
> 5 v 0	$3 \to 2$	$-0.5073(-1.1490, 0.1344)$	$-1.0239(-1.7621, -0.2858)$

Findings

- Both attenuation and strengthening of effects seen
- Substantial attenuation found for the effects of activity on disability
- Consequences may be substantial with regard to inference and conclusions drawn
- Treatment ignored
  - Adjust for marginal distribution of activity as influenced by standard treatment decisions

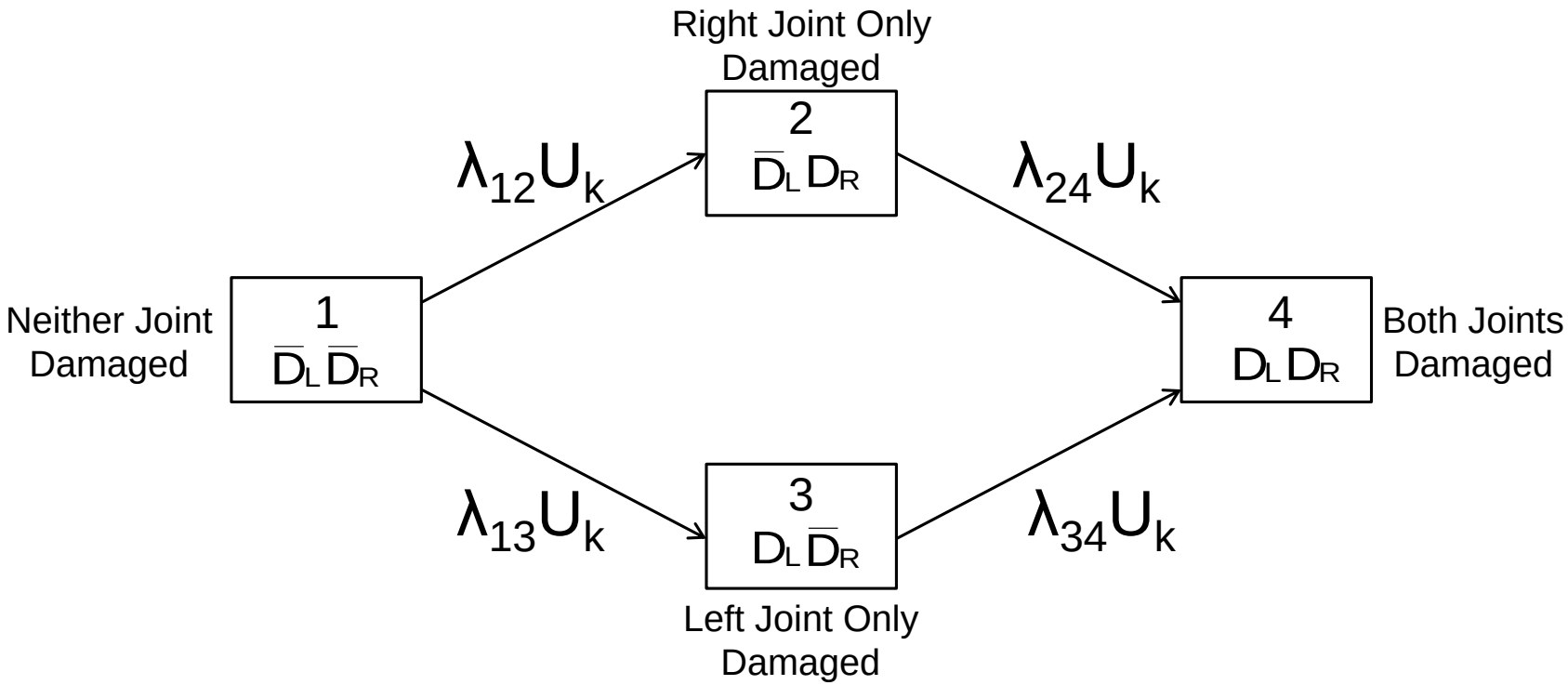
Hand Joints


Distal Interphalangeal Joints (DIP)
Proximal Interphalangeal Joints (PIP)
Metacarpophalangeal Joints (MCP)
PIP1
MCP1

Individual Joint Location Models

- Model damage in the 14 hand joints
- Use Multi-state Models
- Four states for each of the 14 joint locations
  - State 1: Damage in neither hand, $(\bar{D}_L, \bar{D}_R)$
  - State 2: Damage in the right hand only, $(\bar{D}_L, D_R)$
  - State 3: Damage in the left hand only, $(D_L, \bar{D}_R)$
  - State 4: Damage in both hands, (D_L, D_R)
- Patient-specific random effects, U's

Model for Specific Joint Location in Both Hands



$$U_k \sim \Gamma(1/\theta, 1/\theta)$$

Damage and Activity at Joint Level

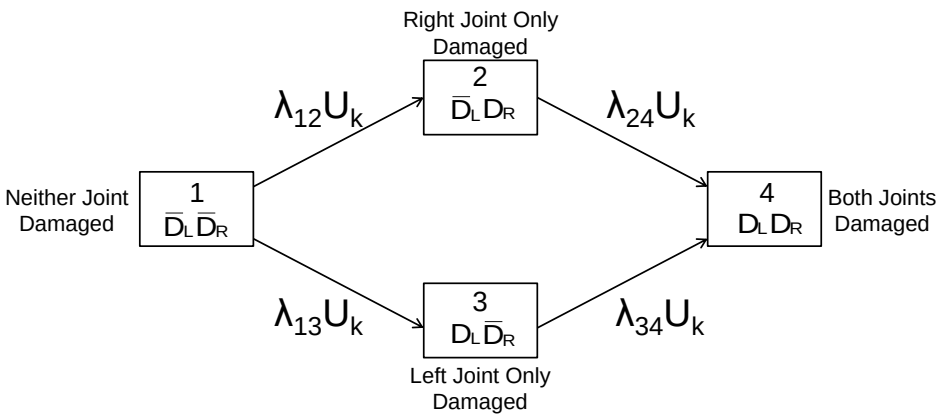
- What is the relationship between damage and (dynamic course of) activity at the individual joint level?
- Define four models (for the lth joint location)

Damage and Activity at Joint Level

- Transition between no damage (State 1) and damage in right hand (State 2)

$$\lambda_{12k}^{(l)}(t) = u_k \lambda_{012} exp(\alpha_{L12} A_L^{(l)}(t) + \tau_{R12} T_R^{(l)}(t) + \epsilon_{R12} E_R^{(l)}(t))$$

  - T_R: Indicator for right joint *tender*
  - E_R: Indicator for right joint *effused (swollen)*
  - A_L: Indicator for left joint *active (tender or swollen)*

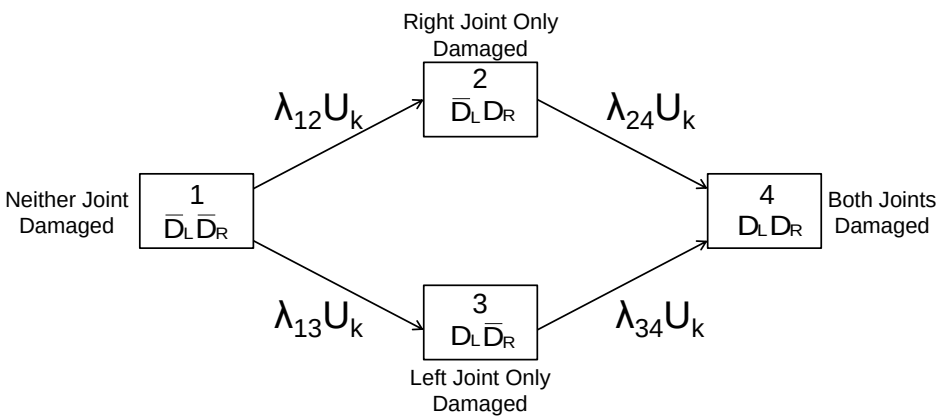


Damage and Activity at Joint Level

- Transition between no damage (State 1) and damage in left hand (State 3)

$$\lambda_{13k}^{(l)}(t) = u_k \lambda_{013} exp(\alpha_{R13} A_R^{(l)}(t) + \tau_{L13} T_L^{(l)}(t) + \epsilon_{L13} E_L^{(l)}(t))$$

  - T_L: Indicator for left joint *tender*
  - E_L: Indicator for left joint *effused (swollen)*
  - A_R: Indicator for right joint *active (tender or swollen)*

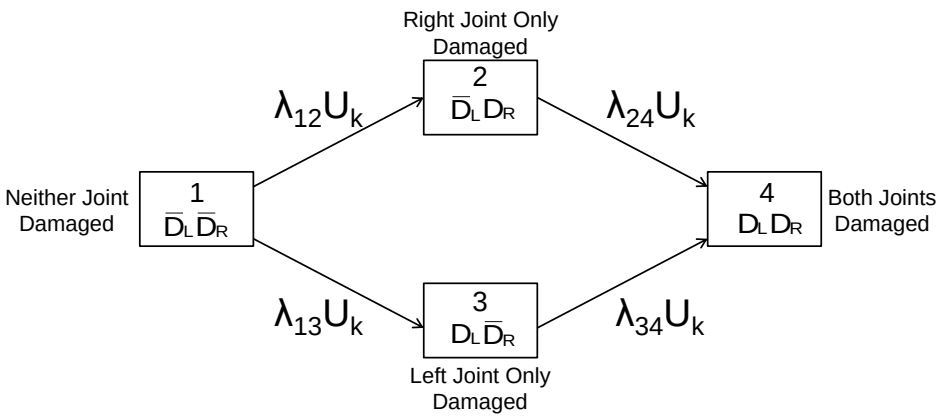


Damage and Activity at Joint Level

- Transition between only right damage (State 2) and damage in both hands (State 4)

$$\lambda_{24k}^{(l)}(t) \quad = \quad u_k \lambda_{024} exp(\tau_{L24} T_L^{(l)}(t) + \epsilon_{L24} E_L^{(l)}(t))$$

  - T_L: Indicator for left joint *tender*
  - E_L: Indicator for left joint *effused (swollen)*

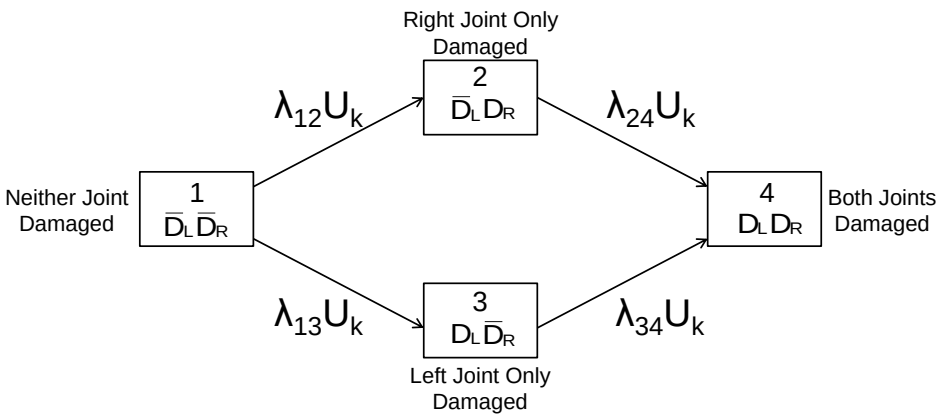


Damage and Activity at Joint Level

- Transition between only left damage (State 3) and damage in both hands (State 4)

$$\lambda_{24k}^{(l)}(t) \quad = \quad u_k \lambda_{024} exp(\tau_{L24} T_L^{(l)}(t) + \epsilon_{L24} E_L^{(l)}(t))$$

  - T_L: Indicator for right joint *tender*
  - E_L: Indicator for right joint *effused (swollen)*



Damage and Activity at Joint Level

- Baseline intensities and parameter effects are constrained to be the same across the 14 hand joint locations
- Because of the panel nature of the data, activity (represented by joint pain and swelling) is assumed to remain constant between clinic visits
- Assume activity variables do not change "state" at the same time damage occurs

Joint Level Activity/Damage Results

Table: Estimated baseline intensities, intensity ratios and random effects variance

Parameter	Estimate	95% confidence interval
Baseline Intensities		
λ_{012}	0.0028	(0.0021, 0.0036)
λ_{013}	0.0027	(0.0021, 0.0034)
λ_{024}	0.0215	(0.0149, 0.0310)
λ_{034}	0.0234	(0.0158, 0.0347)
No previous damage in either joint		
Tenderness in transitive joint	2.76	(2.06, 3.70)
Effusion in transitive joint	4.47	(3.38, 5.90)
Activity in opposite joint	1.18	(0.90, 1.55)
Transitive joint active in past	2.14	(1.68, 2.71)
Opposite joint active in past	1.10	(0.86, 1.41)
Opposite joint damaged		
Tenderness in transitive joint	2.24	(1.51, 3.32)
Effusion in transitive joint	2.19	(1.40, 3.41)
Transitive joint active in past	1.37	(1.00, 1.86)
Random effect variance	3.81	(2.98, 4.88)

Causality

- *Granger causality* is linked to time ordering, and generally specified in discrete time
  - Processes $X(t)$, $Y(t)$ and $U(t)$ represents "all the information in the universe" up to time t.
  - $X(t)$ Granger causal for $Y(t)$ if Y better predicted at time $t+1$ given $U(t)$ than given $U(t) without X(t)$
  - Otherwise, $X(t)$ Granger non-causal for $Y(t)$.
- Local independence (Schweder(1970)) is often taken to infer Granger non-causality in continuous time (Aalen, 1987; Didelez, 2007)
  - Represents a dynamic statistical approach which incorporates time naturally.
  - Natural way to model potential causal relationships.
  - In multi-state models, local dependence/independence is reflected in transition intensities

A. Bradford Hill's Criteria to Infer Causation (1965)

- Non-experimental data show an association.
- Other aspects to be considered.
  1. Strength of the association.
  2. Consistency of the association.
  3. Specificity of the association.
  4. Temporal relationship of the association.
  5. Biological gradient (dose response relationship)
  6. Plausibility
  7. Coherence
  8. Experimental or semi-experimental evidence.
  9. Analogy. (e.g. drugs in pregnancy given thalidomide experience)

US Surgeon's General's Advisory Committee's Report on Smoking and Health (1964)

CHAPTER 3: Criteria for Judgment

Criteria of the Epidemiologic Method

"Statistical methods cannot establish proof of a causal relationship in an association. The causal significance of an association is a matter of judgment ..."

- Consistency of the association.
- Strength of the association.
- Specificity of the association.
- Temporal relationship of the association.
- Coherence of the association.

Interpretation and Causal Implications

- The link between activity and damage appears local/specific to the joint on the particular hand – "local dependence"
- Activity or not in the corresponding joint on the left (right) hand appears *overall* not to influence the damage process on the right (left) hand – "local independence"
- Large effect sizes for these significant associations and probable differential effects of tender only and effused
- Further strengthens argument for a *putative* causal relationship between activity and damage [specificity, strength of association and dose response]

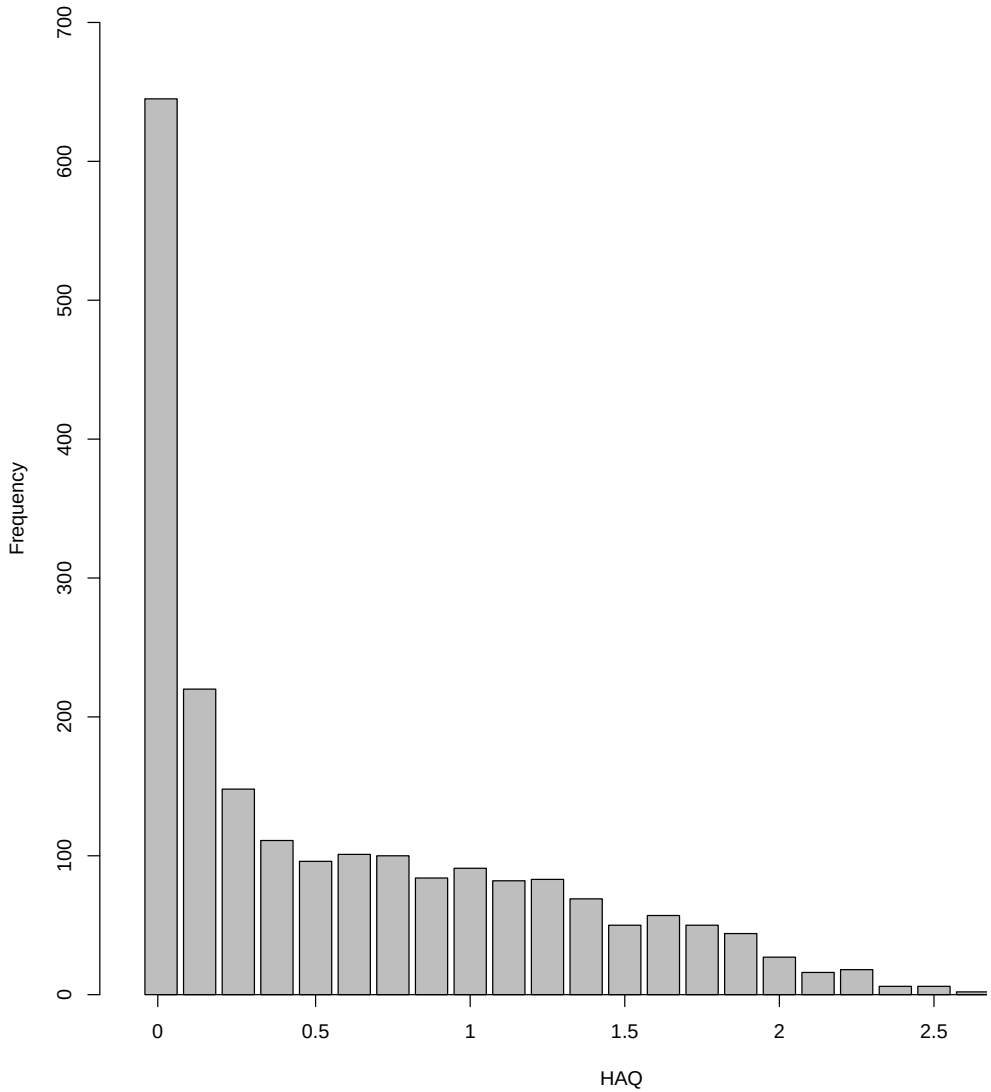
Further Extensions

- Define random effect distributions to be a mixture of zeros and non-zeros to define mover-stayer models.
- Use hidden or partially hidden multi-state models to account for uncertainty in classification.

Motivating Data

- Health Assessment Questionnaire (HAQ): self-report functional disability measure
- Treat HAQ as a continous variable (not categorised as previously)
- HAQ lies in range 0 (no disability) to 3 (completely disabled)
- Data on 382 patients with more than one HAQ measurement. (2107 HAQ observations)

Distribution of HAQ values



Basic Model

- Let Y_{ij} be a semi-continuous variable for the ith $(i = 1, \ldots, N)$ subject at time t_{ij} $(j = 1, \ldots, n_i)$.
- Represented by two variables.
  - The occurrence variable

$$Z_{ij} = \begin{cases} 0 & \text{if } Y_{ij} = 0 \\ 1 & \text{if } Y_{ij} > 0 \end{cases}$$

  - The intensity variable $g(Y_{ij})$ given that $Y_{ij} > 0$.
    - $g(\cdot)$ represents a transformation (e.g., $\log$) that makes $Y_{ij} \mid Y_{ij} > 0$ approximately normally distributed with a subject-time-specific mean.

Regression models

- $$\text{logit}\{\Pr(Z_{ij} = 1)\} = \mathbf{X}_{ij}\boldsymbol{\theta} + U_i,$$
  - $\mathbf{X}_{ij}$ is an explanatory variable vector
  - $\boldsymbol{\theta}$ is a regression coefficient vector
  - U_i is the subject-level random intercept.
- $g(Y_{ij})$ given $Y_{ij} > 0$ follows a linear mixed model

$$g(Y_{ij}) \mid Y_{ij} > 0 = \mathbf{X}^*_{ij}\boldsymbol{\beta} + V_i + \epsilon_{ij},$$

  - $\mathbf{X}^*_{ij}$ is an explanatory variable vector,
  - $\boldsymbol{\beta}$ is a regression coefficient vector,
  - V_i is a subject-level random intercept.
  - $\epsilon_{ij} \sim N(0, \sigma^2_e)$.
- Restrict attention to random intercepts.

Random Effects

$$\begin{bmatrix} U_i \\ V_i \end{bmatrix} \sim \mathsf{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix} \right)$$

- Interpretation of correlation: the presence or absence of disability at one occasion is related to the level of disability, if any, at that and other occasions.
- Variance components, including ρ, usually regarded as nuisance parameters. [$\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ are usual targets of inference.]

Random Effects

$$\begin{bmatrix} U_i \\ V_i \end{bmatrix} \sim \mathsf{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix} \right)$$

- $\rho = 0$ implies separability of the likelihood. Estimation is much easier.
- Incorrect assumption of $\rho = 0$ can lead to bias in estimating the continuous part of the model.
- Problem of informative cluster size: More observations from people more likely to have a non-zero observation, and (usually) more likely to have larger values of Y.

Parallels with nonignorable missingness in a class of "shared parameter models"

- Binary part of two part model $\Leftrightarrow$ logistic random effects model for missing indicators
- Continuous part $\Leftrightarrow$ partially unobserved outcome data
- However,
  - Two part models focus on θ and β.
  - Shared parameter model focusses only on β.

Likelihood

Observed $Y_{ij} = 0$

- Likelihood contribution: $l_{ij} = \{1 - \Pr(Z_{ij} = 1 \mid \boldsymbol{\theta}, u_i)\}$

Observed $Y_{ij} > 0$

- Likelihood contribution:
  $l_{ij} = \{\Pr(Z_{ij} = 1 \mid \boldsymbol{\theta}, u_i)\} \times \left[f\{g(y_{ij}) \mid y_{ij} > 0, \boldsymbol{\beta}, v_i, \sigma_e^2\}\right]$

Likelihood

- $L = \prod_{i=1}^{N} \int_{u_i} \int_{v_i} \prod_{j=1}^{n_i} l_{ij} \times f(u_i, v_i \mid \sigma_u^2, \sigma_v^2, \rho) dv_i du_i$

Back to Motivating PsA Example

- Disease activity is reflected in joints being swollen (*effused*) and/or painful (*tender*)
  - Activity is reversible
- Disease progression is taken to be reflected in damaged joints
  - Damage is generally held to be irreversible
- Interest in whether effects of disease activity and damage on HAQ varies with disease duration.
- PASI (Psoriasis Area and Severity Index) is a disease severity measure for psoriasis.

Parameters (Binary)	**Misspecified model** Estimate (SE)	p	**Full model** Estimate (SE)	p
Intercept	$-1.0199(0.4079)$	0.0129	$-1.0015(0.3746)$	0.0078
Female	$1.9944(0.3603)$	$< .0001$	$2.0080(0.3276)$	$< .0001$
Disease duration	$-0.0027(0.0259)$	0.9169	$0.0156(0.0232)$	0.5207
Active joints	$0.1758(0.0513)$	0.0007	$0.1566(0.0495)$	0.0017
Deformed joints	$-0.0161(0.0321)$	0.6165	$0.0120(0.0260)$	0.6441
PASI score	$0.1941(0.1257)$	0.1233	$0.1754(0.1086)$	0.1071
AJ X Duration	$0.0002(0.0034)$	0.9502	$-0.0003(0.0033)$	0.9403
DJ X Duration	$0.0032(0.0016)$	0.0442	$0.0022(0.0013)$	0.0844
σ_u^2	$4.2519(0.8546)$	$< .0001$	$4.3930(0.8924)$	$< .0001$
ρ	$(\rho = 0)$		$0.9423(0.0373)$	$< .0001$

Parameters (Continuous)	**Misspecified model** Estimate (SE)	p	**Full model** Estimate (SE)	p
Intercept	$0.3176(0.0567)$	$< .0001$	$0.2149(0.0556)$	0.0001
Female	$0.1811(0.0505)$	0.0004	$0.2225(0.0512)$	$< .0001$
Disease duration	$0.0039(0.0033)$	0.2272	$0.0035(0.0032)$	0.2726
Active joints	$0.0219(0.0028)$	$< .0001$	$0.0239(0.0027)$	$< .0001$
Deformed joints	$0.0058(0.0031)$	0.0627	$0.0052(0.0031)$	0.0957
PASI score	$0.0128(0.0140)$	0.3636	$0.0247(0.0134)$	0.0667
AJ X Duration	$-0.0004(0.0002)$	0.0290	$-0.0004(0.0002)$	0.0072
DJ X Duration	$0.0002(0.0001)$	0.1122	$0.0003(0.0001)$	0.0330
σ_v^2	$0.1587(0.0154)$	$< .0001$	$0.1732(0.0166)$	$< .0001$
σ_e^2	$0.0785(0.0040)$	$< .0001$	$0.0774(0.0039)$	$< .0001$
ρ	$(\rho = 0)$		$0.9423(0.0373)$	$< .0001$
$-2\log$ likelihood		2116.0		2018.1
AIC		2178.0		2082.1

Parameters (Continuous)	**Misspecified model** Estimate (SE)	p	**Full model** Estimate (SE)	p
Intercept	$0.3176(0.0567)$	$< .0001$	$0.2149(0.0556)$	0.0001
Female	$0.1811(0.0505)$	0.0004	$0.2225(0.0512)$	$< .0001$
Disease duration	$0.0039(0.0033)$	0.2272	$0.0035(0.0032)$	0.2726
Active joints	$0.0219(0.0028)$	$< .0001$	$0.0239(0.0027)$	$< .0001$
Deformed joints	$0.0058(0.0031)$	0.0627	$0.0052(0.0031)$	0.0957
PASI score	$0.0128(0.0140)$	0.3636	$0.0247(0.0134)$	0.0667
AJ X Duration	$-0.0004(0.0002)$	0.0290	$-0.0004(0.0002)$	0.0072
DJ X Duration	$0.0002(0.0001)$	0.1122	$0.0003(0.0001)$	0.0330
σ_v^2	$0.1587(0.0154)$	$< .0001$	$0.1732(0.0166)$	$< .0001$
σ_e^2	$0.0785(0.0040)$	$< .0001$	$0.0774(0.0039)$	$< .0001$
ρ	$(\rho = 0)$		$0.9423(0.0373)$	$< .0001$
$-2\log$ likelihood		2116.0		2018.1
AIC		2178.0		2082.1

A Marginal Two-part Model

Model that has been considered:

$$\text{logit}\{\text{Pr}(Z_{ij} = 1)\} = \mathbf{X}_{ij}\boldsymbol{\theta} + U_i,$$

$$g(Y_{ij}) \mid Y_{ij} > 0 = \mathbf{X}^*_{ij}\boldsymbol{\beta} + V_i + \epsilon_{ij},$$

$$\begin{bmatrix} U_i \\ V_i \end{bmatrix} \sim \text{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix}\right)$$

A Marginal Two-part Model

Model that will now be considered:

$$\text{logit}\{\Pr(Z_{ij} = 1)\} = \mathbf{X}_{ij}\boldsymbol{\theta} + B_i,$$

$$g(Y_{ij}) \mid Y_{ij} > 0 = \mathbf{X}^*_{ij}\boldsymbol{\beta} + V_i + \epsilon_{ij},$$

where B_i follows the bridge distribution (Wang and Louis, Biometrika, 2003):

$$f_B(b_i \mid \phi) = \frac{1}{2\pi} \frac{\sin(\phi\pi)}{\cosh(\phi b_i) + \cos(\phi\pi)} \quad (-\infty < b_i < \infty)$$

with unknown parameter ϕ $(0 < \phi < 1)$.

Random effects correlation structure

Define the pair of random variables (U_i, V_i) where

$$\begin{bmatrix} U_i \\ V_i \end{bmatrix} \sim \mathsf{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho\sigma_v \\ \rho\sigma_v & \sigma_v^2 \end{bmatrix} \right)$$

And then define B_i by

$$B_i = F_B^{-1}\{\Phi(U_i)\}$$

where

$$F_B^{-1}(x) = \frac{1}{\phi} \log \left[\frac{\sin(\phi\pi x)}{\sin\{\phi\pi(1-x)\}} \right]$$

Advantages of this marginal model

- Can be conveniently implemented in standard software, e.g. SAS NLMIXED.
- Likelihood based so can deal with unbalanced longitudinal data.
- Offers some degree of robustness in estimation of regression parameters when there is departure from the assumed random effects structure.
  - As shown by Haegerty and Kurland (Biometrika, 2001) for GLMMs.
  - Estimation of marginal parameters more robust than subject specific parameters

Particular advantage of bridge distribution

After integration over (B_i, V_i), the marginal probability $\Pr(Z_{ij} = 1)$ relates to linear predictors through the same logit link function as for the corresponding conditional probability.

- If specify marginal structure for binary part as

$$\text{logit}\{\Pr(Z_{ij} = 1)\} = \mathbf{X}_{ij}\boldsymbol{\theta},$$

- then

$$\text{logit}\{\Pr(Z_{ij} = 1 \mid B_i)\} = \mathbf{X}_{ij}\boldsymbol{\theta}/\phi + B_i.$$

Marginal model for HAQ with genetic predictors

- If we look at (time invariant) genetic predictors for HAQ, it is natural to consider marginal effects.

	Binary part				**Continuous part**	
	marginal		conditional		conditional	
	est. (SE)	p	est. (SE*)	p	est. (SE)	p
Intercept	$0.62(0.18)$	0.0005	$1.28(0.37)$	0.0005	$0.46(0.06)$	$< .0001$
B27	$0.47(0.22)$	0.0324	$0.97(0.45)$	0.0325	$0.17(0.08)$	0.0294
DQw3	$-0.22(0.22)$	0.3040	$-0.46(0.45)$	0.3015	$0.1075(0.08)$	0.16
DR7	$-0.48(0.29)$	0.0972	$-0.98(0.59)$	0.0964	$-0.02(0.10)$	0.8775
DQw3:DR7	$0.81(0.38)$	0.0358	$1.66(0.79)$	0.0350	$0.0256(0.13)$	0.85
Age at onset	$0.40(0.09)$	$< .0001$	$0.82(0.18)$	$< .0001$	$0.10(0.03)$	0.0002
Disease duration	$0.19(0.07)$	0.0072	$0.39(0.14)$	0.0067	$0.05(0.02)$	0.0182
Sex (Female)	$1.22(0.19)$	$< .0001$	$2.51(0.41)$	$< .0001$	$0.34(0.06)$	$< .0001$
σ_b^2	$10.64(1.76)$	$< .0001$				
ϕ	$0.49(0.03)$	$< .0001$				
σ_v^2	$0.29(0.03)$	$< .0001$				
σ_e^2	$0.09(0.01)$	$< .0001$				
ρ	$0.98(0.02)$	$< .0001$				

The marginal mean

- It would be convenient if

$$E\{g(Y_{ij}) \mid \mathbf{X}^*_{ij}, Y_{ij} > 0\} = \mathbf{X}^*_{ij}\boldsymbol{\beta}$$

- but, the correct expression is

$$E\{g(Y_{ij}) \mid \mathbf{X}^*_{ij}, Y_{ij} > 0\} = \mathbf{X}^*_{ij}\boldsymbol{\beta} + E(V_i \mid \mathbf{X}^*_{ij}, Y_{ij} > 0).$$

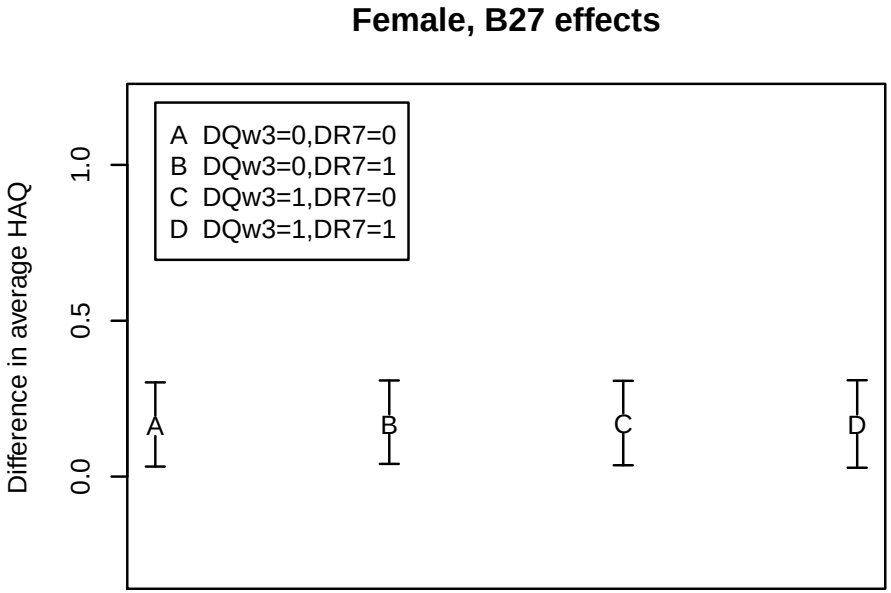
- Two are equivalent only if B_i and V_i are uncorrelated and there are no common explanatory variables in the binary and continuous parts of the model.
- No closed form for $E(V_i \mid \mathbf{X}^*_{ij}, Y_{ij} > 0)$ but can be evaluated numerically..

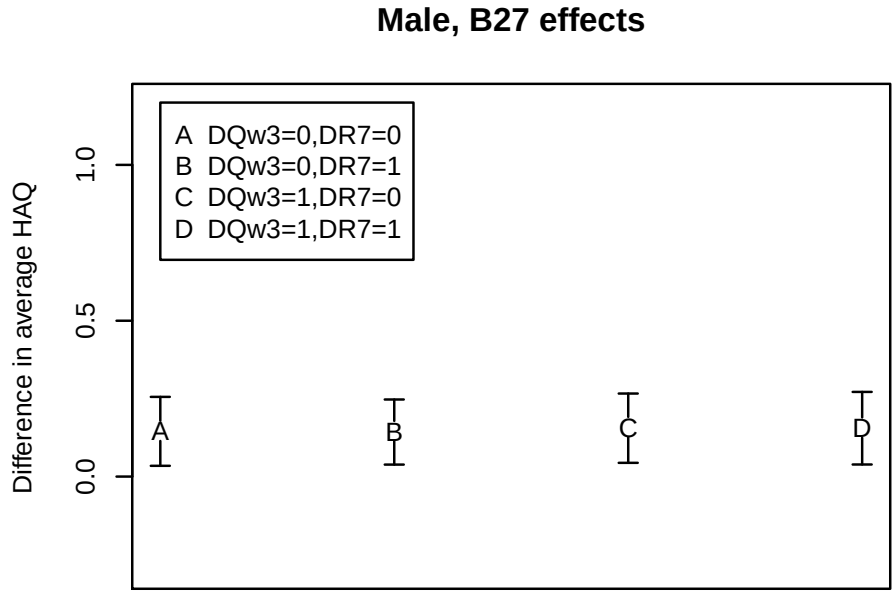
The marginal mean

- Assume $g(.)$ is the identity function for simplicity and suppress dependencies on the explanatory variables.
- Overall marginal mean of the response, $E(Y_{ij}) \equiv E(Y_{ij} \mid \mathbf{X}^*_{ij})$, is given by

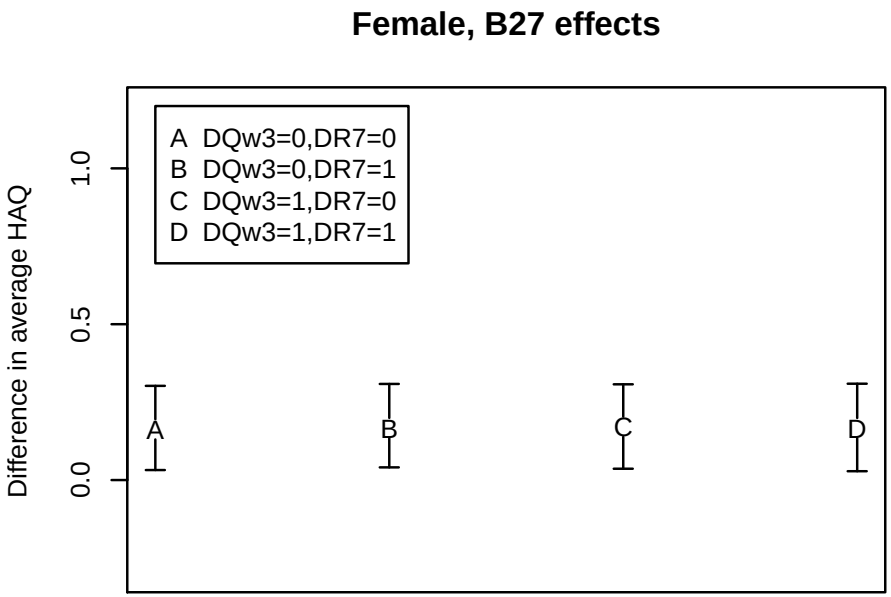
$$E(Y_{ij} \mid Y_{ij} = 0)\mathsf{Pr}(Y_{ij} = 0) + E(Y_{ij} \mid Y_{ij} > 0)\mathsf{Pr}(Y_{ij} > 0)$$

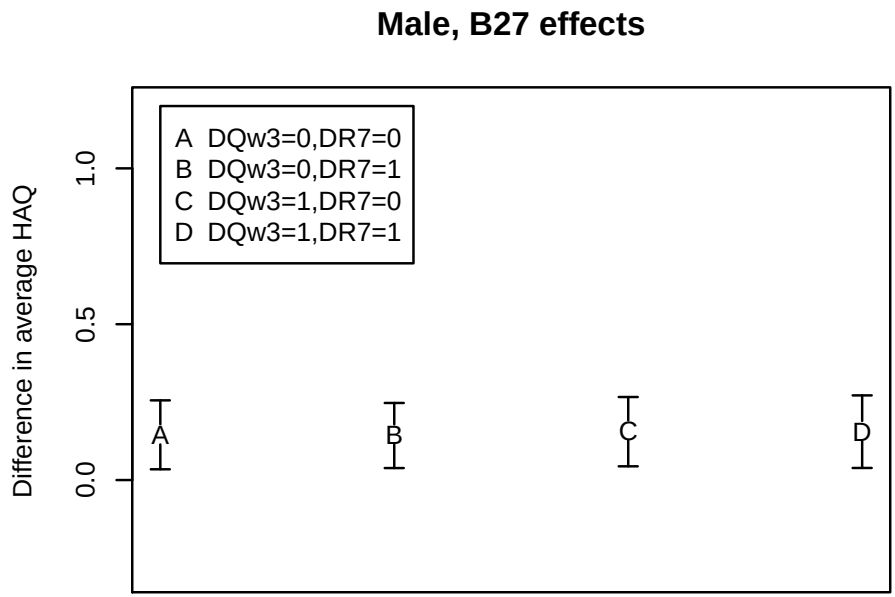
$$= E(Y_{ij} \mid Y_{ij} > 0)\mathsf{Pr}(Y_{ij} > 0).$$



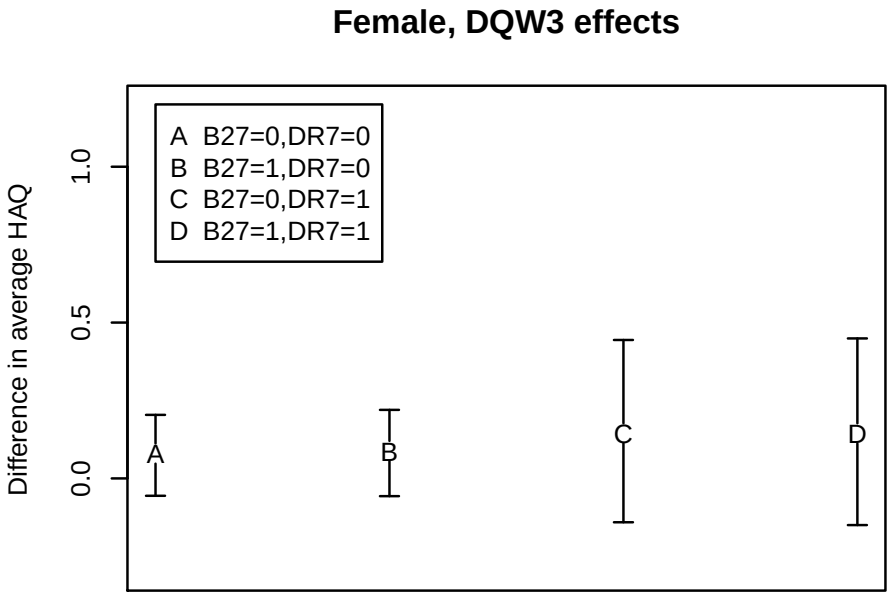


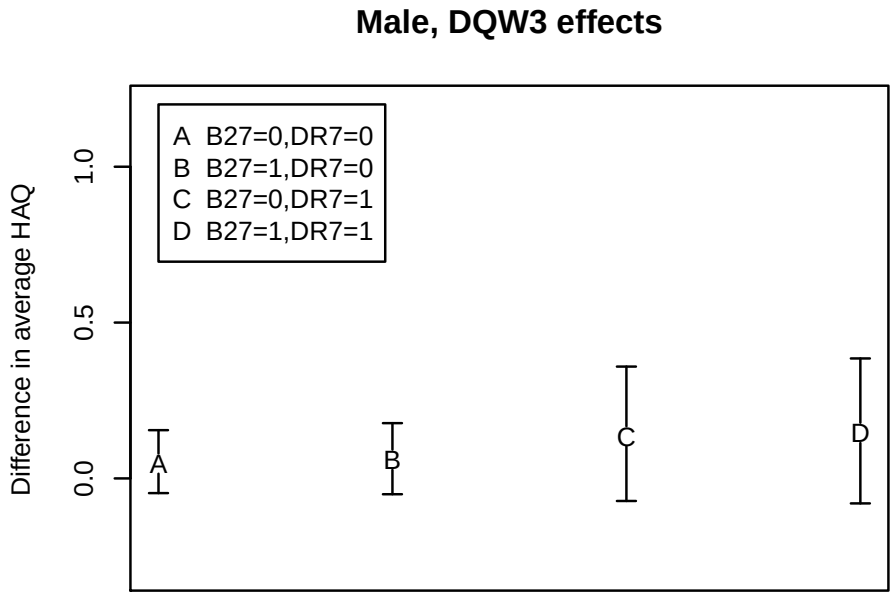
- Calculate contrasts comparing overall expected HAQ with and without specific alleles, stratified by other HLA covariates and holding age at diagnosis fixed at 35 years and disease duration at 15 years.
- Confidence intervals by sampling from the asymptotic distribution of the MLEs of the two-part model.





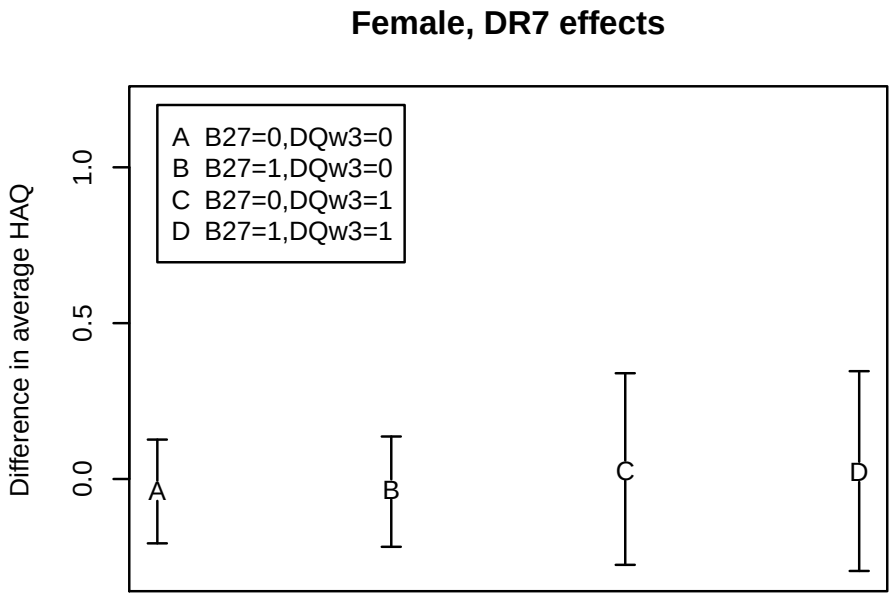
- B27 marginal effects approximately the same across combinations of other variables.
- Confidence intervals exclude zero.

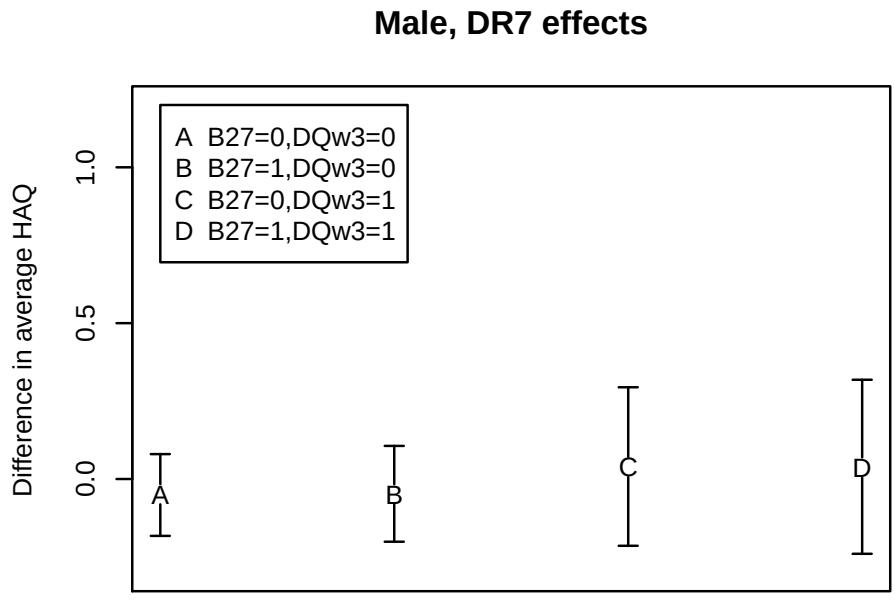




DQW3 X DR7 Interaction (Females)

- D-B: 0.0564 (-0.2062, 0.3232) *DQW3XDR7 effect if B27 present*
- C-A: 0.0648 (-0.1971, 0.3158) *DQW3XDR7 effect if B27 absent*





DQW3 X DR7 Interaction (Females)

- D-B: 0.0564 (-0.2062, 0.3232) *DQW3XDR7 effect if B27 present*
- C-A: 0.0648 (-0.1971, 0.3158) *DQW3XDR7 effect if B27 absent*