Practice of Epidemiology

Statistical Issues in Life Course Epidemiology

Bianca L. De Stavola¹, Dorothea Nitsch¹, Isabel dos Santos Silva¹, Valerie McCormack¹, Rebecca Hardy², Vera Mann¹, Tim J. Cole³, Susan Morton¹, and David A. Leon¹

¹ Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom.
² MRC National Survey of Health and Development, Department of Epidemiology, Royal Free and University College Medical School, London, United Kingdom.
³ Centre for Paediatric Epidemiology and Biostatistics, Institute of Child Health, University College London, London, United Kingdom.

Received for publication September 9, 2004; accepted for publication July 25, 2005.

There is growing recognition that the risk of many diseases in later life, such as type 2 diabetes or breast cancer, is affected by adult as well as early-life variables, including those operating prior to conception and during the prenatal period. Most of these risk factors are correlated because of common biologic and/or social pathways, while some are intrinsically ordered over time. The study of how they jointly influence later (“distal”) disease outcomes is referred to as life course epidemiology. This area of research raises several issues relevant to the current debate on causal inference in epidemiology. The authors give a brief overview of the main analytical and practical problems and consider a range of modeling approaches, their differences determined by the degree with which associations present (or presumed) among the correlated explanatory variables are explicitly acknowledged. Standard multiple regression (i.e., conditional) models are compared with joint models where more than one outcome is specified. Issues arising from measurement error and missing data are addressed. Examples from two cohorts in the United Kingdom are used to illustrate alternative modeling strategies. The authors conclude that more than one analytical approach should be adopted to gain more insight into the underlying mechanisms.

correlation; lifetime; path analysis; regression; structural equation model

Abbreviations: G0, grandparents; G1, study participants; G2, study participants’ offspring.

In the last decade, there has been a growing realization that prenatal and early-life biologic and social factors may play an important role in the etiology of many later-life conditions (1–9), in addition to—or in synergism with—those of well-established adult exposures. This field of inquiry is now referred to as life course epidemiology (10).

Studying complex interrelationships of biologic and social variables over time requires longitudinal information spanning broad periods of life. It also raises analytical problems because temporal, and possibly causal, hierarchies among the exposures need to be taken into account (11, 12). For example, risk factors for breast cancer operate at a number of stages in the life course (figure 1). They also may influence each other; for instance, childhood weight is inversely correlated with timing of puberty (13) and, directly or indirectly, later obesity (14). Age at puberty and adult obesity affect breast cancer risk (15) and thus may mediate the effect of childhood weight. If standard multiple regression models are used to study these variables, the effect of childhood weight would be estimated conditionally...
on all other exposures, thus missing the life course perspective, as we show below. Alternatively, if a joint modeling approach is used, where, for example, age at menarche, adult obesity, and breast cancer risk are all outcomes, their interrelationships would be explicitly estimated, albeit within an assumed multivariate structure.

We discuss these problems by introducing a sequence of increasingly complex models to deal with the temporal and causal hierarchies among the exposure variables. We compare them in terms of interpretability as well as flexibility in dealing with missing data and measurement errors, both of which are common features in life course studies. We illustrate issues and models with two examples: the first studies intergenerational influences on size at birth by using data from a Scottish cohort of children recruited in 1962; the second examines childhood height and its impact on adult leg length with data from a United Kingdom birth cohort of women born in 1946. Our aim is to present a broad analytical framework for studies in life course epidemiology and to highlight its relevance to the current debate on causal modeling (16–18).

MODELS FOR DISTAL OUTCOMES

When a structure among the exposure variables is known or presumed, a distinction can be made between variables that act at the inception (“background”) or in the middle (“intermediate”) of the process that leads to the main (“distal”) outcome of interest (16). In the breast cancer example described above, childhood weight is a background variable, age at menarche and adult obesity intermediate variables, and breast cancer the distal outcome.

Statistical models can offer only simplified representations of reality (19). We classify those relevant in life course epidemiology according to the degree to which they explicitly acknowledge the relations among their components.

Conditional models for one outcome

Typically, standard multivariable regression models define the expectation of the outcome of interest \( Y \), \( E(Y) \) (or a suitable transformation \( g(.) \)), as a function of several explanatory variables, \( X_1, X_2, \ldots, X_k \),

\[
g(E(Y)) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k, \tag{1}\]

where \( g(.) \) is the link function used in generalized linear models (20), and the coefficient \( \beta_j \) \((j = 1, \ldots, k)\) represents the conditional effect attributed to a one-unit increase in \( X_j \) when all of the other variables are held constant. If \( g(.) \) is the identity function and \( Y \) is continuous, equation 1 represents a linear regression model.

A regression model such as equation 1 can be fitted with only background variables, thus avoiding the inclusion of exposures that may be on the causal pathway. This was the approach adopted in the early studies linking birth weight with coronary heart disease (21). Therefore, its estimated coefficients measure the effect of each background variable, controlled for that of the others. The longer the time gap between background variables and distal outcome, however, the greater the possibility of intervening modifying effects (see the debate surrounding the fetal programming hypothesis (1–3, 22–24)).

If all types of exposures (background and intermediate) are included in the same model for \( Y \), the resulting regression coefficients measure mutually adjusted effects, that is, effects of background variables not mediated via the intermediate variables and effects of intermediate variables conditional on the background ones. More specifically, if the background variable \( X_1 \) influences the intermediate variable \( X_k \) through \( X_1 \), \( \beta_j \) would capture only the effect of \( X_1 \) on \( Y \) not mediated by \( X_k \).

A special setting that involves the analysis of background and intermediate exposures arises when repeated measures...
of the same variable are taken over time. In this case, the first available measure acts as background for all the following ones. Consider the influence of childhood anthropometric variables on adult obesity (14). Results obtained after including all childhood measurements in the same regression model for the distal outcome (obesity) are difficult to interpret, especially if the measures are taken close to each other in time, because of their respective conditioning. With two repeated body size measures, say $Z_1$ and $Z_2$, taken at times $t_1 < t_2$, the model

$$g(E(Y)) = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2$$  \hspace{1cm} (2)

could be rewritten as either

$$g(E(Y)) = \beta_0 + (\beta_1 + \beta_2) Z_1 + \beta_2 (Z_2 - Z_1)$$  \hspace{1cm} (3)

or

$$g(E(Y)) = \beta_0 + (\beta_1 + \beta_2) Z_2 - \beta_1 (Z_2 - Z_1),$$  \hspace{1cm} (4)

where the difference $(Z_2 - Z_1)$ represents the change in the explanatory variable, for example, body size, occurring from time $t_1$ to time $t_2$.

Equations 2–4 are different parameterizations of the same model (22, 23). When $Z_2$ is replaced by $(Z_2 - Z_1)$, the conditional effect of $Z_1$ on the transformed dependent variable, $g(E(Y))$, changes from $\beta_1$ to $(\beta_1 + \beta_2)$ (equation 3). A similar switch is observed for the effect of $Z_2$ when it is conditioned on $(Z_2 - Z_1)$ (equation 4). The conditional effect of the difference $(Z_2 - Z_1)$ is either $\beta_2$ or $-\beta_1$, depending on whether you condition on the first or second measure, respectively. So, different interpretations are possible depending on the conditioning variable, as originally discussed by Lucas et al. (22), Horta et al. (25), and Cole (26) and revisited by others (23, 24).

A model with several repeated measures, $Z_1, Z_2, \ldots, Z_K$, taken at times $t_1 < t_2 < \ldots < t_K$, can be reparameterized in terms of the first one, $Z_1$, and all subsequent consecutive increments, $(Z_j - Z_{j-1}), j = 1, \ldots, K$. Thus, equation 3 becomes

$$g(E(Y)) = \beta_0 + \left( \sum_{j=1}^{K} \beta_j \right) Z_1 + \left( \sum_{j=2}^{K} \beta_j \right) (Z_2 - Z_1) + \left( \sum_{j=3}^{K} \beta_j \right) (Z_3 - Z_2) + \cdots + \beta_K (Z_K - Z_{K-1})$$  \hspace{1cm} (5)

and similarly for equation 4. In equation 5, the coefficient for $Z_1$ is the sum of all conditional effects associated with $Z_1, Z_2, Z_3, \ldots, Z_K$; it is the cumulative effect of increasing each $Z_j$ by one unit. Such increases would happen, for example, when a unit change at time $t_1$ has irreversible effects on all following values of $Z$. For instance, one extra centimeter in height at age $t_1$ has an impact on all future height values for a child, since his or her trajectory is shifted upward from then on, holding everything else constant, and therefore has a cumulative impact on the outcome. Similarly, the coefficient for each increment $(Z_j - Z_{j-1}), (j = 2, \ldots, K)$, captures the effect of increasing $Z$ during the $j$th interval, with that change shifting all subsequent $Z$ values.

An alternative parameterization, often used in life course epidemiology (13, 14), replaces the increments with the equivalent changes per unit of time (“velocities”), that is, $(Z_j - Z_{j-1})/ (t_j - t_{j-1})$. In this case, the coefficients for the $j$th period are $(t_j - t_{j-1})$ times the coefficients for $(Z_j - Z_{j-1})$ defined in equation 5.

When the model is not reparameterized but left as in equation 2, a graphical approach—the “life course plot”—may help in interpreting the conditional impact of each repeated measure (26). It involves plotting the regression coefficients against the times when measures were taken (after standardization to make these coefficients comparable). When the coefficients switch sign at some time $t_j$, as in figure 2, there is evidence that changes during $(t_{j-1}, t_j)$ affect the outcome of interest.

### Joint models

Joint models deal with several outcomes simultaneously. In this context, they would explicitly define a presumed process underlying intermediate and distal outcomes. For example, variables may be arranged as shown in figure 3 (top), where $Y$ is the distal outcome and $X_1, X_2$, and $X_3$ the explanatory variables. In this diagram, $X_3$ is assumed to be directly affected by $X_1$ and $X_2$ and thus is an intermediate variable, while $X_1$ and $X_2$ are background variables. Its algebraic equivalent is a system of simultaneous equations, with as many equations as there are intermediate and distal outcomes, which takes the name of path analysis (27):

$$\begin{align*}
  g_1(E(Y)) &= \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \\
  g_2(E(X_3)) &= \alpha_0 + \alpha_1 X_1 + \gamma_2 X_2.
\end{align*}$$  \hspace{1cm} (6)

Here, as before, $E(.)$ stands for expectation, while $g_1(.)$ and $g_2(.)$ are link functions. When $g_1(Y)$ and $g_2(X_3)$ are conditionally normally distributed, estimation can be carried out by maximum likelihood or, avoiding the normality assumptions, by three-stage least squares (28). The indirect
effects of a background variable, say, X₁, on the distal outcome can be computed by multiplying the standardized coefficients (e.g., the effects for a one-standard-deviation increase in X₁) found along each of the paths leading from X₁ to Y and then summing them (18, 29). In this figure, there is only one indirect path from X₁ to Y, via X₃. Thus, if X₁ and X₃ are standardized and the path model is correct (30), the indirect effect of X₁ is α₁ multiplied by β₃.

When Y and X₁ are not normally distributed, penalized maximum likelihood (31, 32) or nonparametric maximum likelihood estimation is used, with the latter recently suggested to improve the estimation properties (33).

When some of the variables are proxies for factors that could not be (or were not) measured precisely, latent variables can be introduced within this framework. For example, several dietary variables may be available via a food frequency questionnaire aiming to measure “usual diet” (34). Thus, each of them is a proxy, or “manifestation,” of an unmeasurable construct that nevertheless is of interest. Similarly, repeated height observations during childhood are manifestations of an underlying growth pattern.

In figure 3 (bottom), three variables, X₁, X₂, and X₃, act as proxy (or “manifest”) measures for the unmeasurable/unmeasured variable U (the convention is to use squares for observed variables and circles for latent variables). The effect of U on Y can be estimated by specifying how the three proxies are linearly related to U and also how U is related to Y:

\[
\begin{align*}
E(X₁) &= μ₁ + λ₁ U \\
E(X₂) &= μ₂ + λ₂ U \\
E(X₃) &= μ₃ + λ₃ U \\
g₃(Y, E(U)) &= β₀ + β₁ U.
\end{align*}
\]

Usually, the latent variable U and its proxy variables are assumed to be normally distributed and the parameters μ₁, μ₂, and μ₃ are set to be zero. The link function for Y, g₃(E(U)), could be of any form, however. Other observed variables could be influencing Y; in this case, the last expression in equation 7 would have an additional term β₄X₄.

The first three expressions in equation 7 define the measurement part of the model because U is not observed but is proxied by X₁, X₂, and X₃. The fourth expression in this equation defines the structural part, that is, in this case, the relation between the unmeasured variable U and the distal outcome Y (35). Together, the measurement and structural models form a structural equation model (29).

Because U is not directly observed and does not usually have a quantifiable metric, its influence on the manifest variables can be measured in terms of an arbitrary metric only. One convention is to use the first of the proxy variables as the reference and thus adopt its metric, for example, that of X₁, so that the effect of the latent variable on Y becomes expressed in terms of X₁ units. Alternatively, the variance of the latent construct is fixed to 1 and its effect on Y estimated in terms of a one-standard-deviation change in the latent variable.

As for path analysis, estimation can be carried out by maximum, penalized maximum (31, 32), or nonparametric maximum likelihood (33), depending on the link functions. Generalizations to include more than one latent variable (and associations among them) are straightforward, although issues of identifiability constrain their numbers (29). Generalizations to discrete latent variables are also within the scope of these models. They involve the concept of latent classes, the probability of belonging to each of them being determined by a higher level latent continuous factor (36, 37), with estimation performed by maximum likelihood with the expectation-maximization algorithm (36, 38, 39). Joint models such as these can be fitted in Mplus (40) and Stata (33) software.

Data quality issues

A major difficulty that arises when analyzing life course studies derives from varying data quality. Because the focus is on different time periods, data from multiple sources (including routine data, e.g., cancer registries) are merged although their variables definition, as well as their
completeness, may vary. The number of subjects for whom data are complete on any subset of the variables of interest can therefore be reduced to a small fraction of the total, while the precision of a variable may depend on when it was collected, for example, because of changes in units (as for birth weight, measured in pounds or kilograms). Thus, measurement errors and missing values affect life course studies to a greater extent than standard observational studies.

Several methods are available to deal with measurement errors (41). For example, when data on two or more proxy variables for an exposure of interest are collected, calibration methods can be used within a two-stage conditional approach (42, 43). Alternatively, as described above, joint models that include latent variables can be fitted (44). When the data are affected by missing values, analyses based on complete records (via conditional or joint models) are rarely appropriate because the incorrect assumption of a “missing completely at random” mechanism would lead to biased results (45). If missingness can be assumed to be “at random,” either of two closely related (46) approaches can be used: 1) imputation methods, with conditional models (47); or 2) maximum likelihood plus the expectation-maximization algorithm (36) or Bayesian simulations (48), with joint models. When, instead, data are suspected to be systematically missing because of unmeasured factors (e.g., informative dropout), extensive sensitivity analyses should be performed, whatever the modeling approach (49–52).

EXAMPLES AND DATA

Two examples arising in the analysis of two United Kingdom cohorts will be used for illustration. The first investigates how maternal and grandmaternal factors influence the size of an offspring at birth; the Children of the 1950s Study is used. This cohort includes all people who, in 1962, participated in a reading survey while attending primary school in Aberdeen, Scotland (53, 54). Data were collected from the participants’ obstetric records and on characteristics of their parents and grandparents. In 1999–2000, 4,497 female study members (78 percent of the total) were anonymously linked to Scottish maternity records leading to birth data on their offspring. Thus, information on three generations (the grandparents, denoted G0; the study participants, G1; and their offspring, G2) can be studied.

The second example focuses on how adult leg length, which has been used in cancer and cardiovascular epidemiology as a marker of childhood environmental factors (55, 56), is determined by different periods of childhood growth. Leg length of participants in the Medical Research Council National Survey of Health and Development was measured by a trained nurse when participants were aged 43 years. The National Survey of Health and Development is a socially stratified birth cohort that includes 2,547 women and 2,815 men born during the week of March 3–9, 1946 (57–59) and followed prospectively, with childhood height measured at ages 2, 4, 6, 7, 11, and between ages 14 and 15 years by trained personnel.

RESULTS

Example 1: intergenerational influences on size at birth

We aimed to investigate how strongly intergenerational factors influence a baby’s size at birth (defined as birth weight standardized for gestational age) by using data from the Children of the 1950s Study (figure 4; (54)). For illustration, we consider only biologic G1 and G0 factors (birth size of the mother, and height and parity of the mother and grandmother) and restrict the analyses to 1,724 first singleton baby girls.

Fitting a multiple linear regression model for G2 birth size on all potential explanatory variables shows that all
Birth size is defined as birth weight standardized for gestational age and standardized to have a standard deviation (SD) = 1, corresponding to about 0.5–0.6 kg in birth weight at most weeks of gestation.

**TABLE 1.** Mutually adjusted coefficients obtained from a conditional regression model for offspring birth size* and from a joint regression model for offspring birth size, maternal birth size, and maternal adult height in the Children of the 1950s Study, (N† = 1,692), Aberdeen, Scotland, 1962–2001‡

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Units/coding</th>
<th>Coef†</th>
<th>95% CI†</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1 birth size</td>
<td>1 SD</td>
<td>0.19</td>
<td>0.15, 0.24</td>
</tr>
<tr>
<td>G1 adult height</td>
<td>1 SD (= 6.0 cm)</td>
<td>0.18</td>
<td>0.12, 0.23</td>
</tr>
<tr>
<td>G1 parity</td>
<td>Parous vs. nulliparous</td>
<td>0.26</td>
<td>0.12, 0.40</td>
</tr>
<tr>
<td>G0† adult height</td>
<td>1 SD (= 5.6 cm)</td>
<td>−0.02</td>
<td>−0.07, 0.03</td>
</tr>
<tr>
<td>G0 parity</td>
<td>Parous vs. nulliparous</td>
<td>−0.06</td>
<td>−0.10, −0.03</td>
</tr>
</tbody>
</table>

**Conditional regression model for G2† birth size**

<table>
<thead>
<tr>
<th>G2† birth size</th>
<th>G1 adult height</th>
<th>G1 birth size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coef</td>
<td>95% CI</td>
<td>Coef</td>
</tr>
<tr>
<td>0.19</td>
<td>0.15, 0.24</td>
<td>0.19</td>
</tr>
</tbody>
</table>

**Joint model for G1 birth size, G1 adult height, and G2 birth size**

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Units/coding</th>
<th>Coef†</th>
<th>95% CI†</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1 birth size</td>
<td>1 SD</td>
<td>0.18</td>
<td>0.12, 0.23</td>
</tr>
<tr>
<td>G1 adult height</td>
<td>1 SD (= 6.0 cm)</td>
<td>0.26</td>
<td>0.12, 0.40</td>
</tr>
<tr>
<td>G1 parity</td>
<td>Parous vs. nulliparous</td>
<td>0.13</td>
<td>0.10, 0.16</td>
</tr>
<tr>
<td>G0 parity</td>
<td>Parous vs. nulliparous</td>
<td>0.03</td>
<td>−0.01, 0.01</td>
</tr>
</tbody>
</table>

* Birth size is defined as birth weight standardized for gestational age and standardized to have a standard deviation (SD) = 1, corresponding to about 0.5–0.6 kg in birth weight at most weeks of gestation.
† N, study size; Coef, estimated regression coefficient; CI, confidence interval; G1, maternal (index); G0, grandmaternal; G2, offspring.
‡ For 32 participants, values were missing for at least one of the explanatory variables used in these models.

By instead fitting the path analysis model equivalent to figure 4, we can deal with these associations simultaneously (table 1, bottom). The estimated direct-effect coefficients corresponding to the arrows leading to the other two joint outcomes (in double boxes), likewise the distal outcome, G2 birth size, show that G1 adult height increases with G0 adult height and G1 birth size but decreases with increasing G0 parity (i.e., it was lower in larger families), while G1 birth size increases with increasing G0 parity and G0 adult height. By multiplying the standardized parameters along the relevant pathways, the indirect effects on G2 birth size of G0 parity, G0 adult height, and G1 birth size can be estimated (table 2). Doing so shows that, although G0 adult height has a negative direct effect (i.e., not mediated), its indirect effect via G1 birth size and G1 adult height is positive (0.13 = 0.49 × 0.18 (via G1 adult height) + 0.20 × 0.19 (via G1

**TABLE 2.** Direct, indirect, and total effects on offspring birth size estimated from the joint model for offspring birth size,* maternal birth size, and maternal adult height in the Children of the 1950s Study, (N† = 1, 692), Aberdeen, Scotland, 1962–2001

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Units/coding</th>
<th>Direct effects</th>
<th>Indirect effects</th>
<th>Total effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Coef†</td>
<td>95% CI†</td>
<td>Coef</td>
</tr>
<tr>
<td>G1† birth size</td>
<td>1 SD</td>
<td>0.19</td>
<td>0.15, 0.24</td>
<td>0.03</td>
</tr>
<tr>
<td>G1 adult height</td>
<td>1 SD (= 6.0 cm)</td>
<td>0.18</td>
<td>0.12, 0.23</td>
<td>0.18</td>
</tr>
<tr>
<td>G1 parity</td>
<td>Parous vs. nulliparous</td>
<td>0.26</td>
<td>0.12, 0.40</td>
<td>0.26</td>
</tr>
<tr>
<td>G0† adult height</td>
<td>1 SD (= 5.6 cm)</td>
<td>−0.02</td>
<td>−0.07, 0.03</td>
<td>0.13</td>
</tr>
<tr>
<td>G0 parity</td>
<td>Parous vs. nulliparous</td>
<td>−0.06</td>
<td>−0.10, −0.03</td>
<td>0.003</td>
</tr>
</tbody>
</table>

* Birth size is defined as birth weight standardized for gestational age and standardized to have a standard deviation (SD) = 1, corresponding to about 0.5–0.6 kg in birth weight at most weeks of gestation.
† N, study size; Coef, estimated regression coefficient; CI, confidence interval; G1, maternal (index); G0, grandmaternal.
TABLE 3. Conditional linear regression models for leg length (cm) at age 43 years on childhood height measures using either absolute childhood height measures or absolute height at age 2 years plus height increments after that age, women in the Medical Research Council National Survey of Health and Development, United Kingdom, 1946–1999

<table>
<thead>
<tr>
<th>Height at age (years)</th>
<th>Coef*</th>
<th>95% CI</th>
<th>Coef*</th>
<th>95% CI</th>
<th>Coef*</th>
<th>95% CI</th>
<th>Coef*</th>
<th>95% CI</th>
<th>Coef*</th>
<th>95% CI</th>
<th>Coef*</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.08</td>
<td>0.03, 0.14</td>
<td>-0.02</td>
<td>-0.08, 0.05</td>
<td>-0.02</td>
<td>-0.08, 0.04</td>
<td>-0.03</td>
<td>-0.09, 0.04</td>
<td>-0.07</td>
<td>-0.14, -0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.37</td>
<td>0.32, 0.43</td>
<td>0.10</td>
<td>0.03, 0.18</td>
<td>0.05</td>
<td>-0.03, 0.13</td>
<td>0.04</td>
<td>-0.05, 0.12</td>
<td>-0.01</td>
<td>-0.10, 0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.38</td>
<td>0.31, 0.46</td>
<td>0.14</td>
<td>0.03, 0.24</td>
<td>0.15</td>
<td>0.04, 0.26</td>
<td>0.08</td>
<td>-0.04, 0.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.31</td>
<td>0.21, 0.41</td>
<td>0.29</td>
<td>0.17, 0.40</td>
<td>0.15</td>
<td>0.03, 0.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.03</td>
<td>0.09, 0.09</td>
<td>-0.10</td>
<td>-0.17, -0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.45</td>
<td>0.38, 0.52</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Height increment between ages (years)

| 2 and 4              | 0.37  | 0.32, 0.43 | 0.49  | 0.43, 0.54 | 0.50  | 0.44, 0.55 | 0.51  | 0.45, 0.57 | 0.56  | 0.50, 0.62 |
| 4 and 6              | 0.38  | 0.31, 0.46 | 0.44  | 0.37, 0.52 | 0.47  | 0.39, 0.55 | 0.57  | 0.49, 0.65 |
| 6 and 7              | 0.31  | 0.21, 0.41 | 0.32  | 0.21, 0.43 | 0.49  | 0.38, 0.60 |
| 7 and 11             | 0.03  | -0.03, 0.09 | 0.35  | 0.27, 0.42 |
| 11 and 15†           |       |         |       |         | 0.45  | 0.38, 0.52 |

* N, study size; Coef, estimated conditional regression coefficient; CI, confidence interval.
† Height was measured between ages 14 and 15 years.

birth size) + 0.20 × 0.19 × 0.18 (via G1 birth size and adult height); table 2), leading to a positive and significant total effect. So, although the conditional model provided an insight into intergenerational direct effects, the joint model led to a more comprehensive summary of their interrelationships.

Example 2: childhood growth and adult leg length

We used data on 2,349 female participants in the National Survey of Health and Development, for whom at least one childhood height or adult leg length measure was available, to investigate which childhood periods are most associated with adult leg length. A series of conditional regression models for adult leg length were fitted by adding each childhood height measure one at a time, starting from age 2 years. Because of missing values, these models are based on different numbers of observations (table 3, top).

At first glance, the results are difficult to interpret because of the changing size and, occasionally, sign of the parameters obtained for the same height measure in different models. The only systematic feature of these models is the consistently larger size and significance of the estimated conditional coefficient corresponding to the oldest age. Model 4 is the exception, however. Here, the oldest age, 11 years, reflects stage of sexual maturation as much as linear growth, and stage of maturation is a poor predictor of adult leg length. Thus, it is not surprising that height at this age has a weak effect on adult leg length when conditioned on earlier stature.

In the model including all available childhood measures (model 5), the conditional effects of height at ages 2 and 11 years are negative, showing that, conditionally on all other childhood measures, being shorter at these ages leads to longer adult leg length. This conclusion is also evident when plotting the equivalent standardized coefficients, as suggested by Cole ((26); figure 5). When heights are replaced by height increments (table 3, bottom), the coefficients in each newly fitted model are the sum of the coefficients in the original one (as shown in equation 5). For example, in model 5, the coefficient for height increments between ages 6 and 7 years is precisely 0.493 = 0.148 − 0.012 + 0.047; that is, it is the sum of the conditional height effects at ages 7, 11, and 15 years. Thus, the coefficients shown in the bottom part of table 3 capture the cumulative effect of a shift in height at one age as it affects a girl’s height at all subsequent ages.

Given the similarity of the conditional coefficients for the earliest height increments, the model can be simplified to include only the intervals 2–7 years, 7–11 years, and 11–15 years (table 4, middle). To obtain comparable measures of effect, the model with yearly height velocities is also reported (table 4, bottom). Here, the estimated conditional
coefficients are multiples of those found when height increments are used; for example, the coefficient for height velocity between ages 2 and 7 years is five times that for the equivalent height difference, 0.55, and $0.55 = 0.19 - 0.10 + 0.45$ is the sum of the corresponding conditional height effects (table 4, top).

An equivalent joint analysis would assume that a girl’s growth profile is determined by a latent process that influences her adult leg length. Thus, we parameterized the growth process in terms of “true” height at age 2 years and height velocities between the ages of 2 and 7, 7 and 11, and 11 and 15 years (figure 6). These are latent variables manifested by the observed heights at 2, 4, 6, 7, 11, and 15 years (the “measurement model”) (figure 7), where the latent variables are equivalent to random coefficients in a generalized linear mixed model (60–62). The structural part of this joint model defines instead how the latent variables influence adult leg length (figure 7).

The measurement and structural models were jointly fitted, the first giving estimated mean growth parameters (table 5), which were consistent with both observed values (table 4, bottom) and standard growth charts (63, 64). The structural model gave estimates of the effects of the growth parameters on adult leg length, which were similar to those found by the conditional model (table 4, bottom). The mainly small numeric differences between conditional and joint results are to be attributed to two main factors. First, the joint model was fitted by using all height measures, not just those at ages 2, 7, 11, and 15 years, with the consequent impact on average height at age 2 years. Second, it dealt with the measurement error inherent in the calculated height velocities by specifying them as latent variables. Thus, the confidence intervals for the joint model parameters are wider, better reflecting the uncertainty about the underlying growth process.

An equivalent joint analysis would assume that a girl’s growth profile is determined by a latent process that influences her adult leg length. Thus, we parameterized the growth process in terms of “true” height at age 2 years and height velocities between the ages of 2 and 7, 7 and 11, and 11 and 15 years (figure 6). These are latent variables manifested by the observed heights at 2, 4, 6, 7, 11, and 15 years (the “measurement model”) (figure 7), where the latent variables are equivalent to random coefficients in a generalized linear mixed model (60–62). The structural part of this joint model defines instead how the latent variables influence adult leg length (figure 7).

The measurement and structural models were jointly fitted, the first giving estimated mean growth parameters (table 5), which were consistent with both observed values (table 4, bottom) and standard growth charts (63, 64). The structural model gave estimates of the effects of the growth parameters on adult leg length, which were similar to those found by the conditional model (table 4, bottom). The mainly small numeric differences between conditional and joint results are to be attributed to two main factors. First, the joint model was fitted by using all height measures, not just those at ages 2, 7, 11, and 15 years, with the consequent impact on average height at age 2 years. Second, it dealt with the measurement error inherent in the calculated height velocities by specifying them as latent variables. Thus, the confidence intervals for the joint model parameters are wider, better reflecting the uncertainty about the underlying growth process.

An equivalent joint analysis would assume that a girl’s growth profile is determined by a latent process that influences her adult leg length. Thus, we parameterized the growth process in terms of “true” height at age 2 years and height velocities between the ages of 2 and 7, 7 and 11, and 11 and 15 years (figure 6). These are latent variables manifested by the observed heights at 2, 4, 6, 7, 11, and 15 years (the “measurement model”) (figure 7), where the latent variables are equivalent to random coefficients in a generalized linear mixed model (60–62). The structural part of this joint model defines instead how the latent variables influence adult leg length (figure 7).

The measurement and structural models were jointly fitted, the first giving estimated mean growth parameters (table 5), which were consistent with both observed values (table 4, bottom) and standard growth charts (63, 64). The structural model gave estimates of the effects of the growth parameters on adult leg length, which were similar to those found by the conditional model (table 4, bottom). The mainly small numeric differences between conditional and joint results are to be attributed to two main factors. First, the joint model was fitted by using all height measures, not just those at ages 2, 7, 11, and 15 years, with the consequent impact on average height at age 2 years. Second, it dealt with the measurement error inherent in the calculated height velocities by specifying them as latent variables. Thus, the confidence intervals for the joint model parameters are wider, better reflecting the uncertainty about the underlying growth process.

An equivalent joint analysis would assume that a girl’s growth profile is determined by a latent process that influences her adult leg length. Thus, we parameterized the growth process in terms of “true” height at age 2 years and height velocities between the ages of 2 and 7, 7 and 11, and 11 and 15 years (figure 6). These are latent variables manifested by the observed heights at 2, 4, 6, 7, 11, and 15 years (the “measurement model”) (figure 7), where the latent variables are equivalent to random coefficients in a generalized linear mixed model (60–62). The structural part of this joint model defines instead how the latent variables influence adult leg length (figure 7).

The measurement and structural models were jointly fitted, the first giving estimated mean growth parameters (table 5), which were consistent with both observed values (table 4, bottom) and standard growth charts (63, 64). The structural model gave estimates of the effects of the growth parameters on adult leg length, which were similar to those found by the conditional model (table 4, bottom). The mainly small numeric differences between conditional and joint results are to be attributed to two main factors. First, the joint model was fitted by using all height measures, not just those at ages 2, 7, 11, and 15 years, with the consequent impact on average height at age 2 years. Second, it dealt with the measurement error inherent in the calculated height velocities by specifying them as latent variables. Thus, the confidence intervals for the joint model parameters are wider, better reflecting the uncertainty about the underlying growth process.

An equivalent joint analysis would assume that a girl’s growth profile is determined by a latent process that influences her adult leg length. Thus, we parameterized the growth process in terms of “true” height at age 2 years and height velocities between the ages of 2 and 7, 7 and 11, and 11 and 15 years (figure 6). These are latent variables manifested by the observed heights at 2, 4, 6, 7, 11, and 15 years (the “measurement model”) (figure 7), where the latent variables are equivalent to random coefficients in a generalized linear mixed model (60–62). The structural part of this joint model defines instead how the latent variables influence adult leg length (figure 7).

The measurement and structural models were jointly fitted, the first giving estimated mean growth parameters (table 5), which were consistent with both observed values (table 4, bottom) and standard growth charts (63, 64). The structural model gave estimates of the effects of the growth parameters on adult leg length, which were similar to those found by the conditional model (table 4, bottom). The mainly small numeric differences between conditional and joint results are to be attributed to two main factors. First, the joint model was fitted by using all height measures, not just those at ages 2, 7, 11, and 15 years, with the consequent impact on average height at age 2 years. Second, it dealt with the measurement error inherent in the calculated height velocities by specifying them as latent variables. Thus, the confidence intervals for the joint model parameters are wider, better reflecting the uncertainty about the underlying growth process.

An equivalent joint analysis would assume that a girl’s growth profile is determined by a latent process that influences her adult leg length. Thus, we parameterized the growth process in terms of “true” height at age 2 years and height velocities between the ages of 2 and 7, 7 and 11, and 11 and 15 years (figure 6). These are latent variables manifested by the observed heights at 2, 4, 6, 7, 11, and 15 years (the “measurement model”) (figure 7), where the latent variables are equivalent to random coefficients in a generalized linear mixed model (60–62). The structural part of this joint model defines instead how the latent variables influence adult leg length (figure 7).

The measurement and structural models were jointly fitted, the first giving estimated mean growth parameters (table 5), which were consistent with both observed values (table 4, bottom) and standard growth charts (63, 64). The structural model gave estimates of the effects of the growth parameters on adult leg length, which were similar to those found by the conditional model (table 4, bottom). The mainly small numeric differences between conditional and joint results are to be attributed to two main factors. First, the joint model was fitted by using all height measures, not just those at ages 2, 7, 11, and 15 years, with the consequent impact on average height at age 2 years. Second, it dealt with the measurement error inherent in the calculated height velocities by specifying them as latent variables. Thus, the confidence intervals for the joint model parameters are wider, better reflecting the uncertainty about the underlying growth process.

An equivalent joint analysis would assume that a girl’s growth profile is determined by a latent process that influences her adult leg length. Thus, we parameterized the growth process in terms of “true” height at age 2 years and height velocities between the ages of 2 and 7, 7 and 11, and 11 and 15 years (figure 6). These are latent variables manifested by the observed heights at 2, 4, 6, 7, 11, and 15 years (the “measurement model”) (figure 7), where the latent variables are equivalent to random coefficients in a generalized linear mixed model (60–62). The structural part of this joint model defines instead how the latent variables influence adult leg length (figure 7).

The measurement and structural models were jointly fitted, the first giving estimated mean growth parameters (table 5), which were consistent with both observed values (table 4, bottom) and standard growth charts (63, 64). The structural model gave estimates of the effects of the growth parameters on adult leg length, which were similar to those found by the conditional model (table 4, bottom). The mainly small numeric differences between conditional and joint results are to be attributed to two main factors. First, the joint model was fitted by using all height measures, not just those at ages 2, 7, 11, and 15 years, with the consequent impact on average height at age 2 years. Second, it dealt with the measurement error inherent in the calculated height velocities by specifying them as latent variables. Thus, the confidence intervals for the joint model parameters are wider, better reflecting the uncertainty about the underlying growth process.
For comparison, the parameters of the joint model were estimated by using the same subset of women contributing to the conditional model ($N = 794$). However, if it is assumed that the missing height and leg length values occurred at random (45), the joint model can be refitted by using data on the 2,349 women for whom at least one height or leg measure was available, via the expectation-maximization algorithm available in Mplus (table 5, right; (40)). The results are marginally different from those found earlier, reflecting the fact that the 794 women in the initial analyses were, on average, taller at younger ages, grew slower from ages 11 to 15 years, and had a shorter leg length than the rest of the women (table 6).

**DISCUSSION**

In this paper, we have described the issues arising when explanatory variables are closely associated because of underlying temporal or biologic processes, a frequent feature of life course studies. Conditional and joint models have been compared by using examples drawn from our work in cardiovascular and cancer epidemiology. These models were chosen for their relative simplicity, the purpose of the analyses being illustrative, so that thorough epidemiologic investigations were restrained and technical details avoided. Other, more complex applications can be found in the life course literature (e.g., Naumova et al. (14), De Stavola et al. (65), and Nitsch et al. (66)). Related joint models can also be found in the survival analysis literature concerned with time-changing internal (i.e., intermediate) exposures, such as CD4 counts in patients with human immunodeficiency virus, as reviewed in Tsiatis and Davidian (67). Issues of measurement error, temporally correlated

---

**FIGURE 6.** Hypothetical piecewise linear model for childhood growth. Velocity was defined as height increment divided by the difference between the ages.

**FIGURE 7.** Simplified path diagram of the relation between height measurements at ages 2, 4, 6, 7, 11, and 15 years and adult leg length. Arrows depicting random variation for each variable are omitted for simplicity. Boxes, proxy variables; circles, latent variables. Velocity was defined as height increment divided by the difference between the ages.
We have used the classification of conditional and joint models to contrast two main analytical approaches. The first class of models is relatively easy to apply but also to misinterpret when the conditioning variables are overlooked. In contrast, joint models may seem ideally suited to deal with life course problems because they explicitly specify the presumed causal and temporal mechanisms for the distal outcome. Further missing data problems can be dealt with directly, if a missing-at-random assumption is appropriate, and measurement error problems by specifying latent variables within a structural equation model. However, several alternative model specifications ("structures") might be appropriate for a particular application. Thus, the choices may be too subjective, especially because formal comparisons are problematic (73).

### TABLE 5. Joint estimation of the measurement and structural models described in figure 7 obtained by using only complete records or all records with at least one of the measures and obtained via EM*-ML* assuming MAR,* women in the Medical Research Council National Survey of Health and Development, United Kingdom, 1946–1999

<table>
<thead>
<tr>
<th></th>
<th>Complete records analysis (N* = 794)</th>
<th>Under the missing-at-random assumption (N = 2,349)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef† Expected mean 95% CI*</td>
<td>Coef Expected mean 95% CI</td>
</tr>
<tr>
<td><strong>Measurement model for latent growth</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height (cm) at 2 years</td>
<td>87.3 87.0, 87.6</td>
<td>87.3 87.0, 87.5</td>
</tr>
<tr>
<td>Height velocity† (cm/year) between ages (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 and 7</td>
<td>6.59 6.53, 6.66</td>
<td>6.51 6.45, 6.57</td>
</tr>
<tr>
<td>7 and 11</td>
<td>5.26 5.19, 5.34</td>
<td>5.27 5.22, 5.32</td>
</tr>
<tr>
<td>11 and 15‡</td>
<td>4.37 4.31, 4.43</td>
<td>4.41 4.36, 4.44</td>
</tr>
</tbody>
</table>

**Structural model for leg length (cm)**

<table>
<thead>
<tr>
<th></th>
<th>Complete records analysis (N* = 794)</th>
<th>Under the missing-at-random assumption (N = 2,349)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef† Expected mean 95% CI*</td>
<td>Coef Expected mean 95% CI</td>
</tr>
<tr>
<td>Height (cm) at 2 years</td>
<td>0.48 0.39, 0.56</td>
<td>0.45 0.38, 0.52</td>
</tr>
<tr>
<td>Height velocity† (cm/year) between ages (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 and 7</td>
<td>2.29 1.67, 2.91</td>
<td>2.49 1.98, 3.00</td>
</tr>
<tr>
<td>7 and 11</td>
<td>1.79 1.34, 2.25</td>
<td>1.96 1.54, 2.31</td>
</tr>
<tr>
<td>11 and 15‡</td>
<td>1.96 1.47, 2.45</td>
<td>2.31 1.84, 2.77</td>
</tr>
</tbody>
</table>

* EM, expectation-maximization algorithm; ML, maximum likelihood; MAR, missing at random; N, study size; Coef, estimated regression coefficient for the joint model; CI, confidence interval.
† Velocity was defined as height increment divided by the difference between the ages.
‡ Height was measured between ages 14 and 15 years.

### TABLE 6. Characteristics of women for whom records on childhood height and adult leg length were or were not complete, Medical Research Council National Survey of Health and Development, United Kingdom, 1946–1999

<table>
<thead>
<tr>
<th></th>
<th>Records complete</th>
<th>Records not complete</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. Mean (SD*)</td>
<td>No. Mean (SD)</td>
</tr>
<tr>
<td>Height (cm) at age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>794 85.08 (4.40)</td>
<td>1,079 84.38 (5.05)</td>
</tr>
<tr>
<td>7</td>
<td>794 119.96 (5.39)</td>
<td>1,199 119.18 (5.88)</td>
</tr>
<tr>
<td>11</td>
<td>794 141.36 (6.99)</td>
<td>1,119 140.49 (7.24)</td>
</tr>
<tr>
<td>15‡</td>
<td>794 158.81 (6.14)</td>
<td>932 158.15 (6.42)</td>
</tr>
<tr>
<td>Height velocity† (cm/years) between ages (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 and 7</td>
<td>794 6.98 (1.00)</td>
<td>839 6.93 (1.08)</td>
</tr>
<tr>
<td>7 and 11</td>
<td>794 5.35 (1.01)</td>
<td>1,013 5.34 (0.98)</td>
</tr>
<tr>
<td>11 and 15‡</td>
<td>794 4.36 (1.13)</td>
<td>841 4.42 (1.14)</td>
</tr>
<tr>
<td>Leg length (cm)</td>
<td>794 75.45 (4.57)</td>
<td>814 75.66 (4.84)</td>
</tr>
</tbody>
</table>

* SD, standard deviation.
† Velocity was defined as height increment divided by the difference between the ages.
‡ Height was measured between ages 14 and 15 years.
Whatever approach is adopted, the main issue is how to deal with life course dependencies while at the same time considering the impact of unmeasured—or poorly measured—factors that influence the pathways of interest. This is not a new topic in epidemiology (74–76), although it has recently become the focus of renewed interest, as demonstrated by the current debates on causal modeling (18, 24, 77–84) and on the role of statistics in causal inference (85–88). Inspired by that debate, we have used the distinction between conditional and joint modeling in the context of life course studies, mirroring that between descriptive and causal modeling (74). Indeed, structural equation models could be viewed as algebraic representations of causal beliefs (18). It must be stressed, however, that either approach is prone to misspecifications and thus should not be singly relied upon for causal inference (16). To achieve robust conclusions, more than one analytical approach should be adopted, with the results compared and inconsistencies investigated, thus carrying out sensitivity analyses in the broader sense (89).

ACKNOWLEDGMENTS

This work was conducted within the framework funded by the Medical Research Council Co-operative Group grant (G9819083) on “Life-course and trans-generational influences on disease risk.”

The authors thank Professors Michael Wadsworth and Diana Kuh for access to the Medical Research Council National Survey of Health and Development data, and Michael Hills and Jonathan Sterne for invaluable comments on an earlier draft of the manuscript.

Conflict of interest: none declared.

REFERENCES