



## Invited Commentary

### Invited Commentary: Beyond Frequencies and Coefficients—Toward Meaningful Descriptions for Life Course Epidemiology

Constance Wang

From the Division of Epidemiology and the Division of Community Health and Human Development, School of Public Health, University of California, Berkeley, CA.

*Received for publication February 17, 2006; accepted for publication February 28, 2006.*

The expected doubling of the elderly population in the United States by year 2030 poses a major challenge to public health and medical care systems because of limited and progressive diminution of resources for public health and medical care for the elderly (1). This challenge highlights the urgent need for effective programs and policies aimed at the prevention of diseases and maintenance of health over the life span of individuals. Such programs must rely on solid, translatable evidence from population-based etiologic studies. However, both etiologic studies of healthy aging and the translation of study findings into programs and policies are exceedingly complex. For example, to identify accurately and precisely the predictors of the onset and the progression of diseases and functional decline, we must contend with not only 1) multiple arrays of time-varying risk and protective factors and 2) multiple morbidities and functional outcomes over time but also 3) synergism among these factors, 4) risk accumulation processes, and 5) latency periods. To develop tailored prevention programs for implementation at optimal times to maximize benefits and cost-effectiveness, decision makers need detailed documentation of the changes in the population distribution of the causes of disease onset and disease progression with increasing age, as well as comprehensive and accurate risk information on priority subgroups.

These challenges for etiologic research and for evidence translation call for inventive approaches in epidemiologic research and practice. One valuable and increasingly popular approach is the life course framework for epidemiologic research and public health program and policy development (2, 3). This approach integrates streams of disciplinary-specific theories and methodologies from biology, human development, public health, and medicine, as well as the social and policy sciences, to account for health and disease

over the life span of individuals. The life course framework unifies theories of disease causation and progression in individuals and populations across many conceptual domains that include 1) multiple multilevel determinants of disease, 2) critical and sensitive exposure time periods, 3) accumulation of risk and protection, and 4) the dynamic interplay among biology, adaptive processes, and sociocultural, environmental, and historical contexts (4).

The life course approach, while interdisciplinary and integrative, presents epidemiologists and other health researchers with alarming complexity. However, it is perhaps the most comprehensive approach epidemiologists and others might take to begin to accumulate answers to the most fundamental of all public health questions: What actions (in the form of intervention programs and public policies) need to be taken at what age across the life span of individuals to prevent disease and enhance health?

Kuh et al. (5) take the life course approach to quantify the effects of physical growth and developmental milestones in childhood, since these affect physical functioning at midlife. To do these analyses, they use data collected from a representative sample of the intended target population, the British cohort born in 1946. More specifically, they used multiple linear regression to quantify the effects on standard physical performance measures at age 53 years of 1) height-adjusted weight, separately for each follow-up age; 2) the effect of each sex-stratified interval weight and height velocity from birth through age 53 years, adjusted for lifetime social class, current physical activity, and current health status; and 3) the residual effects of age at meeting developmental motor milestones, cognitive ability score at age 8 years, motor coordination score at age 15 years, and age at puberty, by increment adjustment for sex, growth trajectories, childhood social class, and other adult risk factors. In doing these analyses,

the authors performed thorough tests for statistical interaction, tests for biologic mediation, and tests for departures from linearity.

On the basis of these analyses, Kuh et al. (5) observed independent positive effects of the following on physical performance at age 53 years: weight gain among boys before age 7 years; meeting motor milestones at the modal age of 12 months; having higher scores on the cognitive test at age 8 years; and having higher scores on the motor coordination test at age 15 years. These findings are valuable, since they identify which antecedents in isolation are most important for better physical functioning at age 53 years, after adjustment for measured confounders. Although this study contributes novel and important information on the cumulative effects of childhood developmental factors on physical functioning in later life, it simultaneously raises many questions. These are fundamental questions implicit in life course epidemiology with important implications for the direction of the field.

### **HOW SHOULD WE EXAMINE INTERRELATED MEASURES OVER THE LIFE SPAN?**

From a developmental point of view, weight and height measures over time or age are correlated intermediate states. These physical characteristics combine with other age-dependent factors in childhood and adolescence to determine body size and body composition in adulthood. Further, an individual's weight and body composition over time or age in adulthood are jointly related to physical activity patterns as well as physical functioning (6). It is possible to conduct analyses to identify which anthropometric measures at what age, independent of all other measures at all other ages, might be related to physical functioning at midlife. However, the inferences from such analyses depend in large part on the handling of sets of time-dependent covariates, confounders, and intermediates (7, 8). The situation is further complicated by the presence of time-varying factors that are simultaneously confounders and intermediate variables (9). For example, in the dynamic relation between weight and physical functioning, a person's physical activity level is simultaneously a time-dependent confounder (related to both weight and physical functioning), as well as a time-dependent intermediate variable (one's body weight influences one's physical activity level, which jointly affect one's physical functioning) (10, 11). The use of such statistical techniques as marginal structural models could yield more robust estimates. From a developmental view, however, information on the relative contribution of discrete weight measures to physical functioning, independent of contextual information, is of questionable value. These independent effects lack essential contextual descriptions: Under what circumstances do certain physical measures at a given age combine with which other age-related factors to lead to better health and functioning in later life?

### **WHAT IS THE CONTRIBUTION OF THE SOCIOECONOMIC CONTEXT?**

From the life course perspective, one's socioeconomic positioning is a time-varying upstream factor that continu-

ously influences the individual's health and functioning over time through multiple downstream intermediate pathways (12, 13). This concept, together with likely residual confounding that arises from the inadequate measurement and the inappropriate categorization of time-dependent socioeconomic positioning (14), makes multiple linear regression with standard adjustments for socioeconomic positioning a particularly problematic model choice. The utility of the study findings by Kuh et al. (5) is limited, since the authors did not treat socioeconomic positioning as an upstream variable, and they did not perform bias level sensitivity analyses (15) to quantify fully and report the uncertainty of their study findings, given likely measurement errors from socioeconomic positioning and other variables. Nevertheless, Kuh et al. do provide valuable descriptions for weight velocities by socioeconomic positioning: Compared with those from the "manual" social class, children from the "non-manual" social class had higher weight velocities through age 7 years. After this age, children from the manual social class had higher weight velocities. Descriptions of this type provide the basis from which investigators may generate additional research questions. For example, among children from the manual social class, what are the predictors of meeting developmental milestones at the modal age? What are the socioeconomic positioning-stratified, age-related dietary changes, especially before and after age 7 years? The answers to this line of inquiry give much needed contextual descriptions to intervention planners and policy makers that could ultimately lead to better decisions.

### **SHOULD WE MOVE AWAY FROM THE ONE RISK FACTOR-ONE OUTCOME APPROACH?**

Most epidemiologic studies related to aging have evaluated sets of risk factors with respect to one disease or functional outcome at a time. Although this approach has yielded useful information on outcome-specific risk factors, there are important knowledge gaps in the fundamental understanding of the risk factor-disease patterns related to aging processes. In particular, morbidities as well as their underlying risk factors (physiologic, psychosocial, and environmental) often occur in clusters, and the relation among and between clusters of risk factors and morbidities may change over time in any given population (16, 17). More attention needs to be devoted to understanding the following: 1) the relations among risk factors and their age-dependent ordering; 2) the underlying structure of the clustering of risk factors and diseases; 3) the causal cascade through which multiple multilevel risk factors and diseases accumulate; and 4) the time-sensitive markers of the disease progression. These domains of information are fundamental to understanding disease causation and progression and, thus, are critical to understanding healthy aging processes (17).

### **HOW DO WE INVESTIGATE DEVELOPMENTAL/AGING PROCESSES AS CUMULATIVE PROCESSES?**

Epidemiologic cohort studies provide rich data sources from which scientists may examine sequential sets of

multilevel determinants of disease as cumulative processes. These data analyses are exceedingly challenging, given that the aim is to tease apart complex data patterns and to estimate higher-order interactions in the presence of background age-period-cohort effects and covariate correlations—multiple collinear variables and variable autocorrelations over time. One approach would be to examine the value added by the application of methods that use entire cascades of life course events (e.g., birth weight, age at smoking initiation, occupation, and age at diagnosis of hypertension) as the predictor variable for functional decline, in comparison with the standard epidemiologic approach of quantifying the relative contribution of each variable one at a time. Another strategy is to look at the value added from the using of the population patterns of complete histories of health states over time (e.g., patterns of transitions to and recovery from disablement over time) as the predictor variable for subsequent morbidities, functional decline, and mortality. The findings from this type of analysis can then be compared with findings from the standard approach on the basis of the presence or absence of disablement as discrete events at different time points as determinants. A more comprehensive model would also include the estimation of multiple interaction effects, as well as the description of the combinations of characteristics that capture as fully as possible the variability across multidimensional domains of variables (18).

The reframing of etiologic research and program and policy development within the life course framework highlights the need to increase our collective capacity to provide comprehensive contextual descriptions while capturing sets of complex causal processes. Such contextual causal analyses are needed, in place of the classic risk factor- or variable-based approaches (19), for investigations of dynamic disease processes, multilevel time-varying mechanisms, and multiple time-dependent interactions. Thus, we need to move beyond the conventional and artificial confines of the decontextualized and nearly fully saturated regression models.

Algorithm-based computational methods (e.g., recursive partitioning) and complex systems models (e.g., dynamic stochastic systems, social network systems) offer the flexibility needed to perform maximally contextual analyses (20–23). Recursive partitioning is a particularly efficient method to perform highly nonlinear predictions, to estimate multiple higher-order interactions, and to detect meaningful data patterns with identifiable error structures (20). With regard to evidence translation, one important unexplored component is the added value of the identification of more complete combinations of variables that define priority subgroups in addition to ethnicity and the social hierarchy variables (17). A related component is the creation of complex stratification variables (17). This would be an alternative to the standard single-variable stratification in risk estimation currently recommended by experts in health disparities research (24, 25). Practical translations of these approaches into strategies to enhance healthy aging will require multidisciplinary research teams with community partners, as well as multiple iterations of theory modification, methods development, application, and evaluation (26).

The relevance of epidemiologic findings hinges on our ability to provide maximally contextual causal inferences.

Standard data tables that present single-variable frequency distributions and beta coefficients from nearly fully saturated models offer woefully unsatisfying data descriptions. We need to embrace theories and methodologies that take full advantage of the richness of life span data and the long view of the life course framework in ways that can move the prevention agenda forward.

---

## ACKNOWLEDGMENTS

Financial support was received from the Robert Wood Johnson Foundation's Health & Society Scholars Program.

The author thanks Len Syme, Ira Tager, and Sara Johnson for helpful comments.

Conflict of interest: none declared.

---

## REFERENCES

1. Smedley BD, Syme SL, eds. Promoting health: intervention strategies from social and behavioral research. Washington, DC: National Academy Press, 2001.
2. Kuh D, Ben-Schlomo Y, eds. A life course approach to chronic disease epidemiology; tracing the origins of ill-health from early to adult life. 2nd ed. Oxford, United Kingdom: Oxford University Press, 2004.
3. Halfon N, Hochstein M. Life course health development: an integrated framework for developing health, policy, and research. *Milbank Q* 2002;80:433–79.
4. Wang C. A call for analytic methods that match theory: analytic suggestions for an interdisciplinary lifecourse approach in epidemiologic research and public health policy. (Abstract). *Am J Epidemiol* 2005;161(suppl):S89.
5. Kuh D, Hardy R, Butterworth S, et al. Developmental origins of midlife physical performance: evidence from a British birth cohort. *Am J Epidemiol* 2006;164:000–00.
6. Sternfeld B, Ngo L, Satariano WA, et al. Association of body composition with physical performance and self-reported functional limitation in elderly men and women. *Am J Epidemiol* 2002;156:110–21.
7. Kaufman JS, Maclehorse RF, Kaufman S. A further critique of the analytic strategy of adjusting for covariates to identify biologic mediation. *Epidemiol Percept Innov* 2004;1:4.
8. De Stavola BL, Nitsch D, dos Santos Silva I, et al. Statistical issues in life course epidemiology. *Am J Epidemiol* 2005;163:84–96.
9. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11:550–60.
10. Tager IB, Haight T, Sternfeld B, et al. Effects of physical activity and body composition on functional limitations in the elderly—application of the marginal structural model. *Epidemiology* 2004;15:479–93.
11. Haight T, Tager I, Sternfeld B, et al. Effects of body composition and leisure-time physical activity on transitions in physical functioning in the elderly. *Am J Epidemiol* 2005;162:607–17.
12. Krieger N. Theories for social epidemiology for the 21st century: an ecosocial perspective. *Int J Epidemiol* 2001;30:668–77.

13. Kaufman JS, Kaufman S. Assessment of structured socioeconomic effects on health. *Epidemiology* 2001;12:157–67.
14. Kaufman JS, Cooper RS, McGee DL. Socioeconomic status and health in blacks and whites: the problem of residual confounding and the resiliency of race. *Epidemiology* 1997;8:621–8.
15. Phillips CV. Quantifying and reporting uncertainty from systematic errors. *Epidemiology* 2003;14:459–66.
16. Berkman L, Singer B, Manton K. Black/white differences in health status and mortality among the elderly. *Demography* 1989;26:661–78.
17. Wang C. Characterizing aging profiles: an application of the grade of membership (GoM) model. (Abstract). *Am J Epidemiol* 2005;161(suppl):S76.
18. Wachter KW. Biosocial opportunities for surveys. In: Finch C, Vaupel J, Kinsella K, eds. *Cells and surveys: should biological measures be included in social science research?* Washington, DC: National Academy Press, 2001.
19. Palloni A, Morenoff JD. Interpreting the paradoxical in Hispanic paradox: demographic and epidemiologic approaches. *Ann N Y Acad Sci* 2001;954:140–74.
20. Zhang H, Singer B. *Recursive partitioning in the health sciences*. New York, NY: Springer Verlag, 1999.
21. Singer BH, Ryff CD, eds. *New horizons in health: an integrative approach*. Washington, DC: National Academy Press, 2001.
22. Philippe P, Mansi O. Nonlinearity in the epidemiology of complex health and disease processes. *Theor Med Bioeth* 1998;19:591–607.
23. Breiman L. Random forests. *Mach Learn* 2001;45:5–32.
24. Jones CP. Invited commentary: “race,” racism, and the practice of epidemiology. *Am J Epidemiol* 2001;154:299–304.
25. Kaufman JS, Cooper RS. Kaufman and Cooper respond to “‘race,’ racism, and the practice of epidemiology.” *Am J Epidemiol* 2001;154:305–6.
26. Syme SL. Social determinants of health: the community as an empowered partner. *Prev Chronic Dis* 2004;1:A02.