Commentary

Commentary on “Disability Trajectories at the End of Life: A Countdown Model” The Problems With Time-to-Death as a Predictor of Disability

Scott M. Lynch

Department of Sociology, Duke University, Durham, North Carolina.

Correspondence should be addressed to Scott M. Lynch, PhD, Department of Sociology, Duke University, Durham, NC 27708. E-mail: scott.lynch@duke.edu.

The study of disability trajectories at the end of life is both interesting and important. Everyone dies, and most adults live well into advanced ages in which they experience some level of physical decline before death. This period of decline is relevant to us as individuals, obviously, but also to society, which ultimately bears much of the financial burden of care via Medicare and Medicaid expenditures. Thus, it is important to understand the disability process in late life. In particular, concepts like “terminal drop” and “terminal decline” are of interest because, if we could recognize who is likely to experience which type of pattern, we might be able to develop interventions to improve quality of death, and we might be able to better predict end-of-life expenditures.

The recent paper by Wolf et al. is intriguing for these reasons. In brief, the authors use time-to-death (TTD) as a predictor of disability trajectories. The key goal is to determine the number and shape of disability trajectories that immediately precede death, ultimately estimating the prevalence of those who experience terminal drop, terminal decline, and other patterns.

Although I think the methodology is interesting and certainly meticulous, there are two key problems with studying disability trajectories as a function of TTD that should make one cautious when employing such an approach. First, treating TTD as a predictor of disability patterns that precede it confuses the direction of causality. Second, the approach suggested by the authors cannot yield any more useful information than a model that does not confuse the direction of causality can.

Putting the Cart Before the Horse

TTD is determined after and, in part, by disability. Using TTD as a predictor of disability reverses this causal order. The authors say that they are proposing a descriptive, rather than causal, model. However, in several places, the authors discuss TTD as a causal variable. For example, in the discussion section, the authors make a causal claim in stating that “…we have shown for the first time that a clear majority of older adults are found in trajectory classes in which time to death is much more influential than age in driving the disablement process.” The use of the words “influential” and “driving” directly imply that TTD has been afforded some causal status. Yet TTD cannot be causal because it comes after its supposed effects.

As another example, the authors suggest early in the paper that TTD can be considered a proxy for unobserved factors like biomarkers. Biomarkers precede disability and death and are very commonly viewed as causal. Thus, the authors argue for treating TTD as a proxy for a causal mechanism, and they argue that, even though there is not a perfect relationship between causally prior variables, like biomarkers, and TTD, including TTD as a predictor “would surely be a substantial improvement over the age-only model typically used.”

The key reason the causal ordering issue is important is perhaps not whether using TTD as a predictor violates philosophical views on causality. To be sure, many statistical models are agnostic with respect to direction of causality. Instead, the key issue is that, because TTD comes at the end of a disability trajectory, it cannot be useful in developing interventions or policy regarding disability. We cannot know what trajectory an individual is on until after s/he dies. So, although TTD may “predict” disability well, the solution is to study the more proximate factors, like biomarkers, and TTD, including TTD as a predictor “would surely be a substantial improvement over the age-only model typically used.”

Of course, a model does not have to have policy or intervention implications in order to have some value. However, here, it is not clear how including TTD necessarily improves over typical models, when TTD itself is unknown and must be estimated for a sizeable proportion of any sample (here, roughly 20%). Furthermore, given that we do not know how these unobserved factors that TTD may represent relate to other observed and unobserved variables, it is not clear how any remaining omitted variable biases may be affected by treating TTD as a proxy. In other words, it is not clear how valuable this approach is, even as a descriptive model, given that we cannot
know the extent of any biases. I discuss this and related issues in greater depth in the next section.

There Is No Such Thing as a Free Lunch

The methodology proposed by the authors is complex, but it must rely heavily on assumptions in order to identify various model parameters. To be sure, all statistical models are assumption-driven. However, some models rely on more and stronger assumptions than others. The fundamental problem with the proposed method is that we know neither the TTD of survivors nor their future disability status, and we simply cannot know these with any degree of confidence from the available data. Instead, assumptions identify both TTD and future disability status and the relationship between the two, and consequently, the resulting disability classes are simply mathematical extrapolations from the assumptions and the observed data. If we change the assumptions, we will change the outcome, and we cannot evaluate the validity of the assumptions in real data. We can evaluate the performance of the method in simulated data, and in the interest of full disclosure, I found that the authors’ method does, in fact, work when the modeling assumptions are met. However, it does not work when the assumptions are violated, and there are many assumptions that may be violated.

In order to highlight the assumptions required, I will briefly discuss the steps involved in the method. First, TTD is modeled using a censored regression model. Second, TTD is simulated for those without a known TTD (i.e., those who survive the study period) using these model results. Third, the simulated, complete set of TTD is entered as a predictor, along with age and other covariates, in a series of quadratic latent class growth models for disability in order to determine the extent to which TTD predicts latent class membership. The second and third steps are repeated in a multiple imputation framework in order to compensate for uncertainty inherent in the imputation of unknown TTD.

Regarding the model for TTD, a large percentage of individuals (~20%) survive the study period. For them, TTD is unknown. A censored regression model assumes a distribution for TTD, and then, based on the assumed distribution for mortality and the known values of TTD as well as the censoring point (i.e., the end of the study period), model parameters, including the mean and variance of the uncensored distribution can be obtained. Here, the authors assume that the log of TTD, conditional on a host of covariates, including observed disability measures and (importantly) age, is a censored normal random variable, with the point of censoring being the end of the study period. This model itself is not problematic; the model is a form of accelerated failure time model, which is just a different parameterization of a more typical hazard model.

The fundamental potential problem with this portion of the modeling strategy is with the distributional assumption for TTD. The imputation method assumes that those with unknown TTD come from the same underlying distribution as those with observed TTD. This assumption is the “missing at random” (MAR) assumption, and we can never know whether this assumption is met (Little & Rubin, 2002). In the current context, I would argue that, regardless of the variables included in the model, there is likely unobserved heterogeneity in mortality that results from biological and other factors for which the authors claim TTD may proxy. Thus, it is unlikely that the MAR assumption for TTD is met, based on the authors’ own argument. More specifically, the rationale for expecting multiple, distinct classes of disability to arise from unobserved biological (and other) factors should also apply to mortality, yet the authors’ model does not allow for mixtures of mortality distributions.

Biases attributable to unobserved heterogeneity are commonly mentioned in mortality literature, and so I do not think this is a major limitation here over what may be found in other research (see Vaupel and Yashin, 1985). My concern is simply that this method relies heavily on the distributional assumption for mortality, whereas the authors simultaneously undermine the argument in explaining why TTD may be useful (as a proxy for unobserved heterogeneity).

The latent class model, given the complete TTD data, is also not inherently problematic by itself. However, consider the data structure and the data available. The data structure for the two, probably most common, latent class modeling packages—proc traj in SAS (used by the authors) and MPlus (used in my subsequent analyses)—is a multivariate (or “wide”) format, meaning that each row in the data set represents a single person, with extra columns for each observed, time-specific disability measure. For those who die, the data are “complete” in the sense that their disability trajectory is fully observed. For those who do not die during the survey period, complete data exist on disability status over the entire study period. Given the wide-data format, disability for survivors is fully observed over the study period, but their mortality is censored. Most importantly, however, their end-of-life disability trajectory is also censored. For decedents, disability is uncensored but must be treated as “missing” for the time-specific measures of disability that occur later in the study period beyond the decedent’s date of death. This necessitates another complication in model estimation: handling of missing disability measures for decedents. Both major packages that estimate latent class models use ML methods to compensate for missing data on the measures forming the trajectories. This approach assumes that the missing data are at least MAR (if not missing completely at random, as the authors of the documentation claim), meaning in the current context that the missing disability measure does not predict missingness. Put another way, the missing disability status for a person missing on the measure is identical to that of another individual with complete data who looks exactly like him/her. The consequence is that trajectories of disability for decedents will be forced to look like trajectories of disability for survivors, and the trajectory classes that emerge will be more heavily weighted toward those of survivors because they are observed on more occasions. Thus, although the authors argue that limiting the analyses to decedents poses a potential selection bias because decedents probably have worse health than survivors, the method for handling missing data for decedents employed by the software the authors use will bias estimates of disability classes/shapes of classes toward that of survivors, before then extrapolating those trajectories mathematically to compensate for the fact that survivors’ actual end-of-life disability measures are unobserved.

Figure 1 illustrates these issues. The figure shows disability trajectories for three hypothetical people who are observed from age 70 until either death or the end of the study period 15 years later (much like the Health and Retirement Study [HRS] used in this paper). All three trajectories are flat, with each person having a probability of about .5 of being disabled across the observation period (note: the slight differences are simply shown for illustration purposes so that all three trajectories are not directly on top of one another). One respondent dies by age 80, and so his TTD is known, as is his complete disability trajectory. Another respondent dies by age 84, and so his TTD and disability trajectory is also completely known. The third individual survives the study period, and so his TTD is unknown, as is his disability status prior to death. As this person’s observed variables
(namely observed disability) are identical to the observed variables for the other two individuals, his TTD will be a function of (a) the decedents’ TTD and (b) the distribution TTD is assumed to follow. Yet, it is unclear that his TTD should come from the same TTD distribution. There is also no clear basis for knowing what the survivor’s unobserved disability (prior to his death) might look like. As the figure shows, his disability may increase, decrease, or remain stable. The other individuals who look like the survivor can only suggest that his disability trajectory will remain flat. The fact that survivors are missing on both disability and TTD means that these unknowns can only be determined by the key assumptions: the distributional assumptions for TTD and disability and the MAR assumption (that survivors will look like decedents). But, these assumptions cannot be tested because there are no individuals for which at least one of the variables—TTD or future disability—is present.

Because of these issues, it is unclear whether it is reasonable to use TTD as a predictor of disability. Nor is it clear that doing so is necessary. To the extent that we already know from prior research (a) that the risk of disability increases with age but is not perfectly predicted by it and (b) that disability predicts mortality and explains part of the association between age and mortality, there is no need to pose the question the authors pose. Instead, we can simply develop a model that simultaneously estimates the number of disability trajectories present in an observed sample and the degree to which each predicts mortality. This approach would yield the same conclusions as those obtained by the authors, does not violate the rules of causal order, and does not involve the difficulties the method developed by the authors does.

An Alternative Approach

I illustrate such an approach using data similar to that used by the authors. Specifically, I use data from the 1998, 2000, 2002, 2004, 2006, 2008, and 2010 waves of the HRS. I limit the data to persons who were 70–74 years of age—a single 5-year birth cohort. Given the multivariate format for the data, it is difficult to treat age as a time-varying variable (indeed, it is unclear how age was handled by the authors in their analyses; entering only age at baseline confounds cohort differences and age-related change). I used MPlus version 7.3 (Muthen and Muthen, 1998–2012) for the analyses. I included indicators for the presence of any activities of daily living (ADL) limitation at each wave (just as the authors did), with sex, race (black, other vs. white), and education (years of schooling) as predictors of class membership. MPlus allows the estimation of a censored regression model for an outcome of latent class membership. Thus, I also included TTD as a censored normal variable predicted by disability class.

I estimated latent class growth models with linear and quadratic growth factors for 1–5 classes. Comparison of Bayesian information criterion (BIC) statistics and substantive results indicated a three class model fit the data best. Figure 2 shows the trajectories of probabilities of ADL limitation from age 70 on for each class. Members of Class 1, consisting of just under half of the population (43%) have a stably high probability of being nondisabled across the observed time period. This low probability of being disabled is coupled with a long life expectancy—approximately 14 years.

Members of Class 2 start the study period with a relatively high probability of being disabled (~.6) and experience a slow increase in risk of disability across age. Members of this class can expect to live 6.2 years on average, as indicated by the asterisk on the trajectory in the figure.

Members of Class 3 start the study period with a low risk of being disabled, roughly comparable to that of members of Class 1. However, they experience a relatively rapid decline in ability before the expected end of life at age 75.5. For those who survive beyond age 75, the probability of remaining nondisabled falls from around .8 to around .2 by age 80. Thus, this class could be considered to experience “terminal drop,” when compared with the slower, long-term and stable decline of Class 2.

In terms of the covariates, men are more likely than women to be in Class 3—the class that experiences rapid decline with relatively short life expectancy—that Class 1—the class that experiences
stable nondisability—but less likely than women to be in Class 2 (vs. Class 1), the class that experiences persistently high risk of disability. Blacks and persons of other races are more likely than whites to be in either Class 2 or Class 3 (vs. Class 1). Finally, those with more education are more likely than those with less education to be in Class 1 than either Class 2 or Class 3.

All in all, these analyses are not inconsistent with those found by Wolf et al. However, because these analyses do not reverse causal ordering, they provide results that are more useful than those obtained using TTD as a predictor. Here, I found that those in classes with high or increasing levels of disability had substantially shorter life expectancies than those in the nondisabled class. I also found that roughly a quarter of the population appears to experience a terminal decline before death, while just under a third appears to experience a terminal drop, based on the shape of the disability trajectories around mean life expectancy for each class. Half of the population—those in Class 1—experience virtually no disability, and their life expectancy is outside the observation period. We can use this information to predict mortality risk, and even timing, based on already-observed disability status. For those who are nondisabled, we can predict that their mortality is probably not imminent—at least it is certainly not as imminent as someone with high levels of disability—but we cannot know what their future disability pattern will look like.

References

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). New York: John Wiley.

Muthen, L. K., & Muthen, B. O. (1998–2012). *MPlus user’s guide* (7th ed.). Los Angeles, CA: Muthen & Muthen.

Vaupel, J. W., & Yashin, A. I. (1985). Heterogeneity’s ruses: some surprising effects of selection on population dynamics. *The American Statistician*, 39, 176–185.