Letter to the Editor

RE: "QUANTILE REGRESSION—OPPORTUNITIES AND CHALLENGES FROM A USER'S PERSPECTIVE"

We agree with Beyerlein's argument that quantile regression is an underutilized method in epidemiology [1]. We have found it particularly helpful in our own research when modeling body mass index (BMI) (weight (kg)/height (m)^2) as an outcome variable [2, 3]. For example, we compared linear, logistic, and quantile regression results for associations of maternal and childhood characteristics with adult body size in women and found different inferences for each statistical model [2]. Specifically, maternal weight gain during pregnancy was associated with body size in the upper quantiles for adult daughters, whereas maternal BMI was associated with body size in the lower quantiles for adult daughters [2]. In contrast, rapid growth in childhood was associated with all quantiles of adult body size, and the inferences were robust to statistical model [2]. We also observed similar findings in sibling analyses using quantile regression [3]. The merits of the quantile approach, particularly for models of BMI, were also outlined in an editorial [4] that accompanied our paper [2].

Analyses using fixed cutoff values for the outcome variable are certainly easy to interpret, especially when the fixed cutoffs are universally accepted diagnosis standards (e.g., clinical cut points of blood pressure to define hypertension). There are also applications in which quantiles and percentiles provide natural interpretations. For example, pediatric growth charts [5, 6] consist of age-dependent quantile/percentile curves and are widely used to screen childhood growth. In general, quantile regression offers a systematic examination of the covariate influence on the entire response distribution. This additional information can often provide a much richer interpretation of the data. For example, in their analysis based on quantile models, Terry et al. [3] estimated and compared the distributions of BMI at 7 years of age between different early-life growth patterns while simultaneously adjusting for prenatal factors. Such comparisons provide useful insights, particularly on the effect of exposures on the extreme of BMI, that could be missed if one relies solely on mean-based analysis or analysis using a fixed cut-off value. Moreover, it is often the case that high-risk subgroups often have either high or low values for risk factors, so when modeling intermediate risk factors that are continuous, such as BMI, it can be very useful for interpretation to use quantile regression. For example, in a recent study, Yang et al. [7] found that the fat mass and obesity–associated gene (*FTO*) was associated not only with the mean of BMI but also with the variance. Thus, the influence of *FTO* on BMI differs based on the variance, and its impact could be examined through a quantile-based approach.

In the statistical literature in recent years, the value of quantile regression has received great attention. More recently, the approach has been applied to research questions in genetics [8]. As a result, there has been a rapid growth in methodology development for quantile regression. For example, rapid-estimation algorithms and variable-selection methods have been developed for high-dimensional data [9–11]. In addition, estimation algorithms have been adapted for longitudinal data [12, 13] for better efficiency. Methods are also now available to handle missing data in quantile regression applications [14], as well as errors in measurement [15] and censored outcomes [16].

In summary, quantile regression has many applications that are useful to epidemiologists. With recent methodological advancement in quantile regression, it is a great time to consider this modeling tool for epidemiology applications.

REFERENCES

1. Beyerlein A. Quantile regression—opportunities and challenges from a user's perspective. *Am J Epidemiol*. 2014;180(3):330–331.
2. Terry MB, Wei Y, Esserman D. Maternal, birth, and early-life influences on adult body size in women. *Am J Epidemiol*. 2007; 166(1):5–13.
3. Terry MB, Wei Y, Esserman D, et al. Pre- and postnatal determinants of childhood body size: cohort and sibling analyses. *J Dev Orig Health Dis*. 2011;2(2):99–111.
4. Gillman MW, Kleinman K. Antecedents of obesity—analysis, interpretation, and use of longitudinal data. *Am J Epidemiol*. 2007;166(1):14–16.
5. Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. *Stat Med*. 1992;11(10): 1305–1319.
6. Wei Y, He X. Conditional growth charts. *Ann Stat*. 2006;34(5): 2069–2097.
7. Yang J, Loos RJ, Powell JE, et al. *FTO* genotype is associated with phenotypic variability of body mass index. *Nature*. 2012; 490(7419):267–272.
8. Warrington NM, Wu YY, Pennell CE, et al. Modelling BMI trajectories in children for genetic association studies. *PLoS One*. 2013;8(1):e53897.
9. Portnoy S, Koenker R. The Gaussian hare and the Laplacian tortoise: computation of squared-error versus absolute-error estimators. *Stat Sci*. 1997;12(4):279–300.
10. Belloni A, Chernozhukov V. ℓ_1-penalized quantile regression in high-dimensional sparse models. *Ann Stat*. 2011;39(1):82–130.
11. Wu Y, Liu Y. Variable selection in quantile regression. *Stat Sin*. 2009;19:801–817.
12. Koenker R. Quantile regression for longitudinal data. *J Multivar Anal*. 2004;91(1):74–89.
13. Geraci M, Bottai M. Linear quantile mixed models. *Stat Comput*. 2014;24(3):461–479.

14. Wei Y, Carroll RJ. Quantile regression with measurement error. *J Am Stat Assoc*. 2009;104(487):1129–1143.
15. Wei Y, Ma Y, Carroll RJ. Multiple imputation in quantile regression. *Biometrika*. 2012;99(2):423–438.

Ying Wei[1] and Mary Beth Terry[2]
(e-mail: ying.wei@columbia.edu)
[1] *Department of Biostatistics, Columbia University, New York, NY*
[2] *Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY*

Editor's note: *In accordance with* Journal *policy, Beyerlein was asked if he wished to respond to this letter, but he chose not to do so.*