Supplemental Digital Content

eText 1. Concepts of Causality

The following is an explanation of Hill's criteria of causation. An association accompanies the strength of the finding. An association presents the consistency, or *generalizability* between one exposure variable and an event. An association shows the specificity and temporality, which means one exposure variable needs to occur before an event happens. An association demonstrates the *gradient* or dose-response relationship between an exposure variable and outcome. An association reflects the plausibility based upon knowledge of each exposure and outcome variable. An association depicts the coherence of the relationship among several other experimental designs, and potentially, through an analogy of similar relationships. Rothman's sufficient-component cause model [1] highlights that there may be a number of *sufficient causes* for a given disease or outcome, but a *component cause* that must be present in every sufficient cause of a given outcome is referred to as a *necessary cause*.

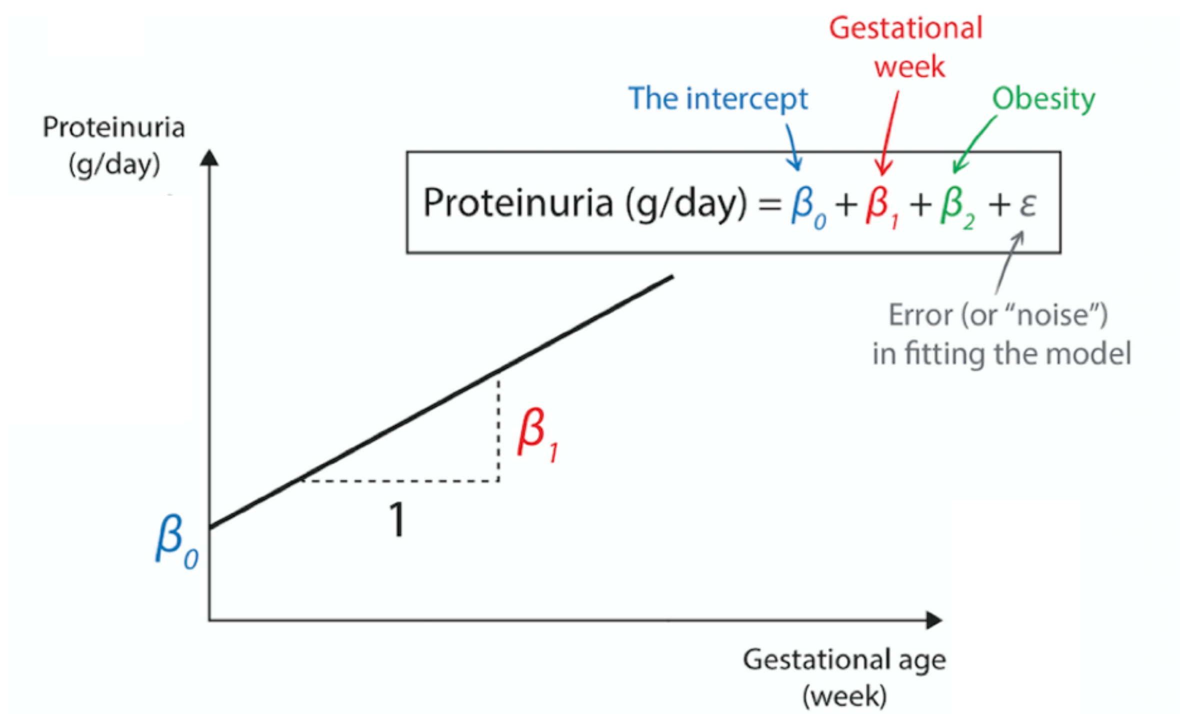
Another useful concept when considering confounding and *causal inference* is "counterfactual theory."[2–4] Counterfactual conditions occur when necessary antecedents are in false or counter to the facts. Counterfactual theory reasons that an event A *causally depends* on B if, and only if, (1) if B had occurred, then A would have occurred, and (2) if B had not occurred, then A would not have occurred [5]. In anesthesia, one example might be the delivery of a depolarizing neuromuscular blocker, succinylcholine, to the case scenario's pregnant patient with a potentially difficult airway. If the patient cannot be intubated or ventilated in another manner, the neuromuscular paralysis would result in her death, which would not have occurred without delivery of the succinylcholine. However, there are also other factors at play such as weight gain, reduced functional residual capacity, a higher risk of aspiration, airway edema, etc. These factors make the relationship between only two variables overly simplistic. Therefore, clinical understanding and analytic and statistical tools are needed to validly estimate the relationships between exposures and outcomes.

eText 2. Key assumptions of linear regression models

Linear regression consists of a number of baseline assumptions that influence the inferences made for the parameters estimated. It is essential to ask a few key questions about these baseline assumptions before developing or interpreting regression models. For example, when developing or interpreting regression models, one needs to consider essential assumptions as follows [6–8]. First, a linear relationship exists between the independent and dependent variables. Other relationships might be dealt with, for example by using a non-linear regression. Second, the data and residuals are normally distributed like a bell curve. The residual refers to the amount of variability in a dependent variable that remains after accounting for the variability due to the independent variables. Third, there is *homoscedasticity* that the dependent variable exhibits similar amounts of variance across the range of values for an independent variable. Last, the observations such as measurement of systolic BP need to be *independent* of each other. The non-independence of the measurements will result in non-independence in the errors of measurement. Because the influence of the variables themselves is often not independent, a multiple linear regression would be performed.

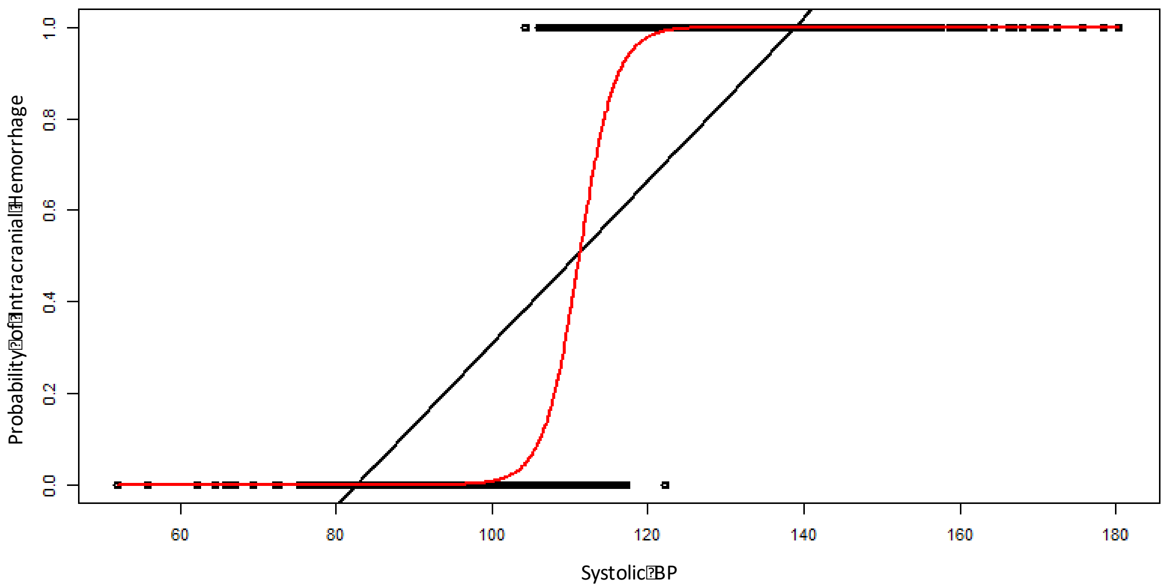
eFigure 1. Proteinuria and gestational age

Figure legend: β's are parameters, with β_0 representing the model intercept (the usual baseline degree of proteinuria that may exist) and β_1 representing the model slope. In this case, β_1 (also called the β coefficient or parameter estimate of gestational age) represents the mean change in the outcome, in this case change protein in the urine (for example, grams per day) for each unit change in gestation (for example, per week). ε represents the error or noise in fitting the model (that is, a variable to capture all other factors that affect proteinuria other than gestational age), which is reasonable to assume does not vary with the change in gestational age.



eFigure 2. Intracranial hemorrhage and systolic blood pressure

Figure legend: The “odds” of the event happening (intracranial hemorrhage in this case) is the probability of the event divided by the probability of no event (p / [1 – p]). Using odds rather than probabilities removes the upper range of an estimate, but not the lower range of an estimate; therefore one can derive an estimate of the association that is higher than an ‘equal’ odds (>1) or lower than an ‘equal’ odds (<1), but has a lower range at 0. By applying the natural log to the odds, one removes the lower range and thus transforms the probability from (0,1) to (-∞,+∞), which is the fitted red curve for the logistic regression with systolic BP in the model.



eReferences.

1. Rothman KJ, Greenland S: Causation and causal inference in epidemiology. Am J Public Health 2005; 95:144–50
2. Hernán MA: A definition of causal effect for epidemiological research. J Epidemiol Community Health 2004; 58:265–71
3. Hernán MA: Does water kill? A call for less casual causal inferences. Ann Epidemiol 2016; 26:674–80
4. Glass TA, Goodman SN, Hernán MA, Samet JM: Causal inference in public health. Annu Rev Public Health 2013; 34:61–75
5. Lewis D: Causation. J Philos 1973; 70:556
6. Slinker BK, Glantz S a.: Multiple linear regression: Accounting for multiple simultaneous determinants of a continuous dependent variable. Circulation 2008; 117:1732–7
7. Austin PC, Goel V, Walraven C Van: An introduction to multilevel regression models. Can J Public Heal 2001; 92:150–4
8. Austin PC, Tu J V, Alter DA: Comparing hierarchical modeling with traditional logistic regression analysis among patients hospitalized with acute myocardial infarction: Should we be analyzing cardiovascular outcomes data differently? Am Heart J 2003; 145:27–35