fact that the two-tailed test is more stringent and more conservative (46,47).

VIII. *P* VALUE

The *P* value is used in a procedure called *hypothesis testing*. *P*, which stands for probability, can be any number between 0.0 and 1.0. According to Whitley and Ball (48), "values close to 0 indicate that the observed difference is unlikely to be due to chance, whereas a *P* value close to 1 suggests there is no difference between groups other than that due to random variation." According to Motulsky (49), "*P* value is simply a probability that answers the following question: If the null hypothesis were true … what is the probability that random sampling … would result in a difference as big or bigger than the one observed?"

Whitley and Ball (50) explain why hypothesis testing is needed, using the example of a drug (nitrate) for preventing deaths from heart disease. Even if there is no real effect of nitrate on mortality, sampling variation makes it unlikely that exactly the same proportion of patients in each group will die. Thus, any observed difference between the two groups may be due to the treatment or it may simply be due to chance. The aim of hypothesis testing is to establish which of these two explanations, treatment versus chance, is more likely.

Hypothesis testing can involve asking whether the mean of a sample has a statistically significant difference from the mean of a population (51). The question of whether there is any difference takes the form of the "null hypothesis." The null hypothesis is that there is no statistically significant difference between the sample and the population.

Where a clinical study involves comparing a study drug group and a placebo group (or study drug sample group and an entire population), the null hypothesis is that there is no statistically significant difference between the study drug group and the placebo group (or no difference between the study drug sample group and the entire population).

The sample mean from the study drug group and the population mean (or the sample mean from the study drug group and the sample mean from the placebo group) can be used to calculate a *P* value. This *P* value is then applied to the null hypothesis.

In hypothesis testing that involves the *null hypothesis*, the researcher does not ask, "Does the study drug work better than the placebo?"

The *null hypothesis* only asks, "Does the study drug work the same as the placebo?"

The question asked by the null hypothesis is the more conservative of these two questions. According to a number of authors (52,53) the null hypothesis is a "straw man" hypothesis.

[46]Norman GR, Streiner DL. Biostatistics. 3rd ed. Hamilton, Ontario: B.C. Decker, Inc.; 2008. p. 56.

[47]Motulsky H. Intuitive biostatistics: A nonmathematical guide to statistical thinking. 2nd ed. New York, NY: Oxford University Press; 2010. p. 99.

[48]Whitley E, Ball J. Statistics review 3: hypothesis testing and *P* values. Critical Care. 2002;6:222–25.

[49]Motulsky H. Intuitive biostatistics: A nonmathematical guide to statistical thinking. 2nd ed. New York, NY: Oxford University Press; 2010. p. 104.

[50]Whitley E, Ball J. Statistics review 3: hypothesis testing and *P* values. Critical Care. 2002;6:222–25.

[51]Jones D. Pharmaceutical statistics. Pharmaceutical Press: Chicago, IL; 2002. pp. 154–6.

[52]Durham TA, Turner JR. Introduction to statistics in pharmaceutical clinical trials. PhP Pharmaceutical Press, Chicago, IL; 2008. p. 76.

[53]Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. Designing Clinical Research, 3rd ed. Lippincott, Williams, and Wilkins, New York, NY, 2006. p. 58.