Reference

[1] Kernan WN, Viscoli CM, Makuch RW, Brass LM, Horwitz RI. Stratified randomization for clinical trials. J Clin Epidemiol 1999;52: 19–26.

An Error in An Old Paper Illustrates the Need for Data/Code Archives - Author response

Dear Drs. Knottnerus and Tugwell,

Thank you for inviting our reply to the letter by Dr. McCullough regarding our review of stratified randomization for clinical trials, which was published almost 20 years ago in JCE [1]. We are delighted that this article is still finding an audience but chagrined that Dr. McCullough has found an error in Table 1.

Unfortunately, we no longer have drafts of our article or the original programming used to calculate the probabilities. In response to Dr. McCullough, however, we recreated the calculations in the table using SAS v.9.4. After generating test data sets with the appropriate factor distributions (15% and 30%) and total numbers of patients, we used the random number function (RANUNI) to sort the data and assign half of the patients to each treatment group sequentially. Note that for a factor prevalence of 15%, we assumed that 5 of 30 patients and 8 patients of 50 patients had the factor. The results of the new simulation study over 10,000 trials are shown below.

The findings of the new study are all within 2 percentage points of the original findings except when the factor prevalence was 30% and number of patients was below 100. In those cases, we found 56% probability (instead of the reported 43%) and 45% probability (instead of 38%) for trials of size 30 and 50, respectively.

We are not able to determine the reason for these discrepancies. We note that Dr. McCullough suggests that there may have been a problem with the original random number generator and we agree this is possible. We did not reference the software used to create the original calculations and recognize this as an error.

Dr. McCullough cites our article in recommending a policy for archiving data and codes for submitted publications. We agree with the principle that data should generally be available for most studies, but our article involved no original data collection and only one calculation. The errors in Table 1 do not change any of our conclusions or recommendations. In 20 years, no other reader has found an error in our article or requested information on our calculation. Archiving our code would probably have served little purpose that could not have been better served had we simply referenced our calculation method.

Table 1

(Recreated). Probability that simple randomization will result in 2 treatment groups that differ by more than 10% for the proportion of patients with a prognostic factor according to the prevalence of the factor in the whole cohort

Total no. of patients in the trial	Probability of 10% difference according to prevalence of the prognostic factor in the entire cohort	
	15%	30%
30	33%	56%
50	24%	45%
100	12%	28%
200	4%	11%
400	0.5%	3%

Walter N. Kernan*
Professor of Medicine
Yale School of Medicine
New Haven, CT

Catherine M. Viscoli
Research Scientist
Yale School of Medicine
New Haven, CT

Ralph I. Horwitz
Professor of Medicine
Temple University School of Medicine
Philadelphia, PA
*Corresponding author. Professor of Medicine, Department of Internal Medicine, Yale School of Medicine, Room 407A, Harkness Hall A, New Haven, CT 06510. Tel.: +1-203-785-5208; fax: +1-203-737-3306.
E-mail address: walter.kernan@yale.edu (W.N. Kernan)

Reference

[1] Kernan WN, Viscoli CM, Makuch RM, Brass LM, Horwitz RI. Stratified randomization for clinical trials. J Clin Epidemiol 1999;52:19–26.