[4] Tedeschi F. A clarification on odds ratios as conditional risk ratios. J Clin Epidemiol 2018;97:122−3.
[5] Redelmeier DA, Tibshirani RJ. A simple method for analyzing matched designs with double controls: McNemar's test can be extended. J Clin Epidemiol 2017;81:51−5.

https://doi.org/10.1016/j.jclinepi.2018.11.009

"An error in an old paper illustrates the need for data/code archives"

To motivate their much-cited article, Kernan et al [1] presented their Table 1, herein reproduced. The purpose of the table is to illustrate the probability of covariate imbalance when drawing two samples from a common distribution in which a specific prognostic factor has a specified prevalence. For example, if a sample of 50 observations has 15 with the specified prognostic factor (let us call them "hits" for short) and it is divided into two subsamples of 25 each, it would be most desirable for subsample A to have eight hits and subsample B to have seven (or conversely); no closer balance can be obtained. In this situation, the percentage of hits in subsample A is 32% and the percentage of hits in subsample B is 28%, for a difference of 4%. As the text of the Kernan article makes clear, Table 1 is concerned with the probability that the difference will be more than 10%; a difference of exactly 10% is not considered extreme. Five numbers in the table are wrong, which are denoted by asterisks.

The number of hits in subsample A has a hypergeometric distribution. R code for computing the relevant probabilities is given by:

```
nobs < − 50 # total number of observations
hits < − 15 # number of desired items
myseq < − seq (0, hits)
dhyper (myseq, m = hits, n = nobs − hits,
k = nobs/2).
```

Once one recognizes that this is a hypergeometric problem, it is immediately obvious that the numbers 33% and 24% in Table 1 are problematic. Fifteen percent of 30 is 4.5 and 15% of 50 is 7.5. The hypergeometric distribution is not defined for noninteger numbers.

Let us consider the case when the total number of observations is 50 and the percentage of hits is 30%. In this case, there are 15 hits to be divided between the two subsamples. Only two possible allocations result in an absolute difference less or equal to 0.10: seven hits in subsample A (and therefore eight in subsample B), 7/25−8/25 = −0.04, which occurs with probability 0.231; or eight hits in subsample A (and seven in subsample B), 8/25 - 7/25 = 0.04, which also has a probability of 0.231. Therefore, the overall probability that the difference exceeds 0.10 is 1 - 0.462 = 0.538, markedly different from the claimed probability of 0.38. Similarly, it can be shown that 10% should be 9% and 27% should be 28%.

Table 1
Probability that simple randomization will result in two treatment groups that differ by more than 10% for the proportion of patients with a prognostic factor according the prevalence of the factor in the whole cohort[a]

Total no. of patients in the trial	Probability of 10% difference according to the prevalence of the prognostic factor in the entire study cohort	
	15%	30%
30	33%*	43%
50	24%*	38%*
100	10%*	27%*
200	3%	9%
400	0.3%	2%

Asterisk denotes incorrect number.
[a] Calculated in a simulation study involving 10,000 hypothetical trials.

When contacted by email for the simulation code, the corresponding author replied that he no longer had it. This is reasonable, given that the article is 20 years old. However, lost forever is knowledge of how the probabilities in Table 1 were calculated. Fortunately, the errors in this table do not affect the conclusions of the article. However, we have no assurance that such is the case for as-yet-undiscovered errors in other articles. The published article is just part of the scholarship and research that leads to the conclusions stated in the article. Behind that lies the data and code, which is just as crucial to understanding how results were obtained and enabling other researchers to build on what was published. If the article is really just part of the scholarship, why does not the journal make all the scholarship available to other researchers and require that both data *and* code be available? The replication literature has demonstrated repeatedly that data alone are insufficient to ensure reproducibility; the code must be available, too. There were no data for the Kernan article, but there certainly was code. If data have to be uploaded, it takes no additional effort or resources to also require that code be archived, too.

B.D. McCullough
Department of Decision Sciences & MIS
Drexel University
Philadelphia, PA 19104, USA
Department of Decision Sciences & MIS, Drexel University, GHall 742, Philadelphia, PA 19104, USA.
Tel.: 215 895 2134; Fax: 215 895 2907.
E-mail address: bdmccullough@drexel.edu

DOI of original article: https://doi.org/10.1016/S0895-4356(98)00138-3.
Conflict of interest statement: None.

Reference

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

https://doi.org/10.1016/j.jclinepi.2018.11.012

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.

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%

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.

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.

https://doi.org/10.1016/j.jclinepi.2018.11.011

DOI of original article: https://doi.org/10.1016/j.jclinepi.2018.11.012.
No funding was provided to support this reply.
This serves as our statement that we have no conflict of interest related to our submission.