Medicine & Science Journal 2023;1(1):e202332
https://doi.org/10.53150/msj.v1i1.32

BRIEF ARTICLE

The role of sex in sample size determination

J.M.T.A. Meessen* & S. Garattini*

* Institute for Pharmacological Research Mario Negri IRCCS.

Abstract

The sample size calculation in randomized clinical trials mainly depends on the incidence of the outcome measure in the control group. Sample size calculation should also take into account the differences in outcome rates between male vs. female genders. This issue is, however, often not considered, leading to over- or under-estimation of the outcome distribution and ultimately to underpowered trials with erroneous conclusions. Hence, this short article discusses examples related gender and sample size and provides indications for an optimal estimation.

Keywords: simple size, sex, gender

The size of your sample of patients for a randomized clinical trial on the effect of a treatment on an outcome measure depends on the incidence of the outcome measure in the control group. Therefore, you must determine the precise distribution of your outcome measure in the control population. Once you know this (e.g. 50% of the patients die), you can make an educated guess of the effect of the new treatment on this distribution - for instance mortality will be reduced to 40%). Currently, sample size calculations do not take into account the differences in outcome rates between the sexes, and calculate on the basis of the overall average of the population. This over- or under-estimation of the outcome distribution may lead to an underpowered trial, giving erroneous conclusions. Thus, it is inherently wrong to calculate sample sizes by just assuming males and females are similar.

An example: Atrial fibrillation (AF) is a common cardiac arrhythmia and is associated with a five-fold increase in risk of stroke. The increased risk depends on various stroke risk factors. Despite a higher reported prevalence of AF in males [1], several studies have described a higher risk for stroke in women than in men, especially in those aged 75 years or older. [2] The overall prevalence of

Received: February 16, 2023
Revised: February 16, 2023
Accepted: March 01, 2023
Published: March 14, 2023

Article information are listed at the end of this article.

AF with hospitalization in Lombardy was 2.4% between 2002 and 2013, the prevalence rising with age (0.39% <65 years and 8.45% ≥75 years). AF with hospitalization was more common in males (2.67%) than females (2.15%), p<0.001. Females were consistently older and had fewer risk factors for AF, such as hypertension and diabetes mellitus. A meta-analysis including 993,600 patients found a significantly higher risk of stroke in female AF patients (HR: 1.24, 95%CI: 1.14-1.36, p<0.001). [3]

This finding was confirmed by *Marzona et al.*; the cumulative risk for stroke was higher for females with AF, and on correcting for age, chronic HF, hypertension, diabetes mellitus, prior stroke, myocardial infarction, peripheral artery disease, chronic kidney disease, oral anticoagulant drugs and antiplatelet drugs, this increase in risk remained significant (HR: 1.18, 95CI: 1.14-1.21). [4]

If a RCT were to assess the effects of a certain treatment on the incidence of stroke in patients with AF, it would have to take account of the differences in the proportion of male or female patients who experience a stroke over the course of the trial. If we ignore the differences between sexes we could take the proportion of patients who experience a stroke as a whole: 20100/315383 - an incidence of 6.4%. [4] Using these data to estimate a sample size for a hypothetical trial in which we believe that our new treatment leads to a 25% relative risk reduction (RRR) of stroke, including standard parameter values (80% power and alpha 0.05), we would need to include 6282 patients (Table 1).

TABLE 1 - Sample size for stroke in patients with atrial fibrillation.

	TT control group	RRR	TT experimental group	Effect size	Power	α	N group	N total
All	6.4%	25%	4.780%	0.071	80%	0.05	3141	6282
Males	5.3%	25%	3.975%	0.063	80%	0.05	3935	7870
Females	7.4%	25%	5.550%	0.075	80%	0.05	2764	5528

Sample size calculation for categorical endpoint, proportion of stroke in AF patients was taken from *(Marzona et al., 2020)*. Analyses done using the pwr-package, pwr.2p.test-command in RStudio.

If someone ran the trial with 6282 patients, assuming equal distribution of the sexes, we would end up with 3141 males and 3141 females, of whom half are randomized to one treatment arm and half to the other. Any subgroup analyses for the effect of the treatment by sex on these small data sets would therefore be underpowered (Table 2).

TABLE 2 - Power calculation for stroke in patients with atrial fibrillation.

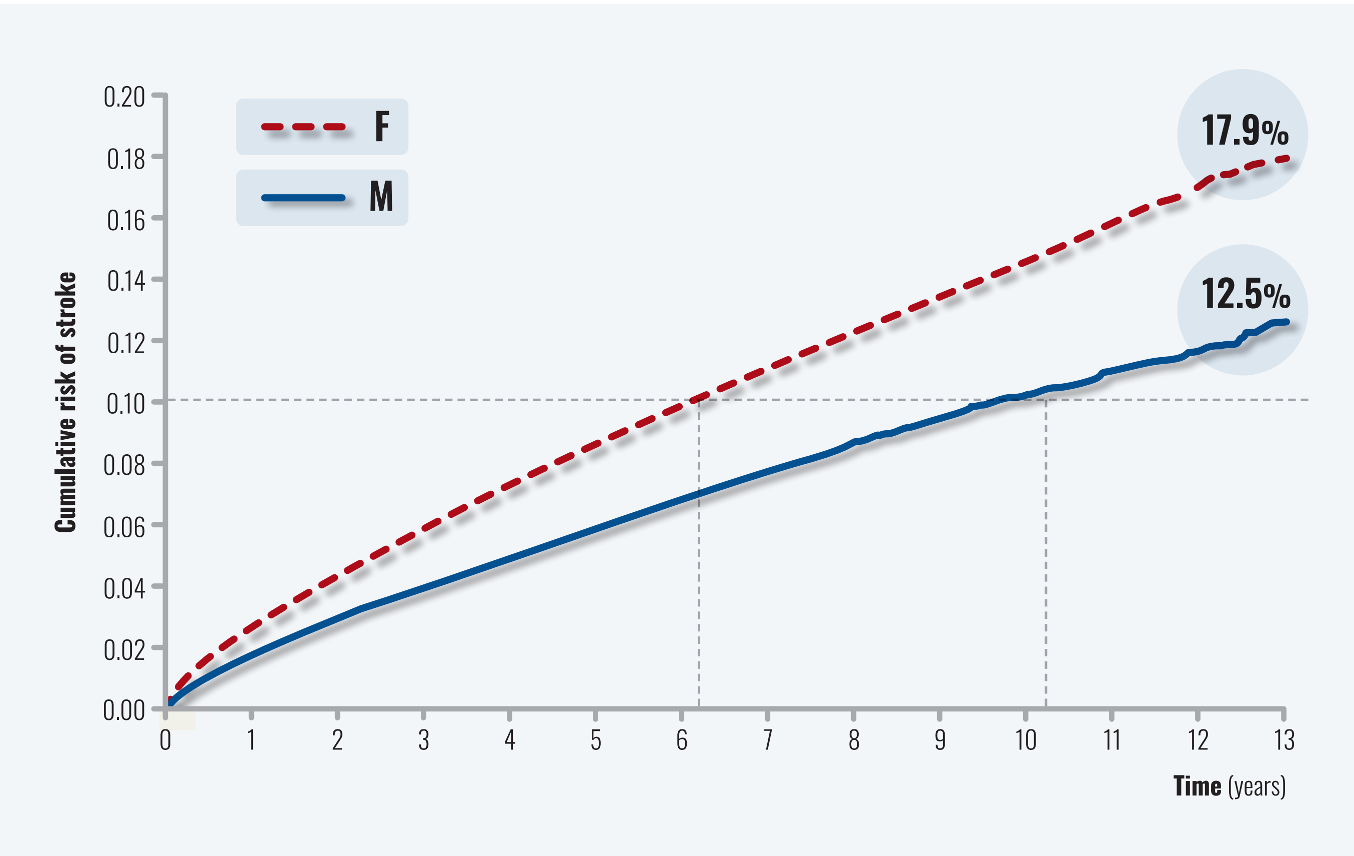
	N RCT	N group	π control group	RRR	TT experimental group	Effect size	α	Power
Males	3141	1570	5.3%	25%	3.975%	0.063	0.05	42.5%
Females	3141	1570	7.4%	25%	5.550%	0.075	0.05	56.0%

Analyses done using the pwr-package, pwr.2p.test-command in RStudio.

Here we calculated the power, assuming that the RRR would be 25% of our treatment. For males, the power is only 42.5%, meaning that the chances of a Type 2 error, accepting a false null hypothesis, is 57.5% (1-power). However, as we saw in the studies by *Marzona et al., 2018 and Marzona et al., 2020*, the proportions of AF patients who experience a stroke in the first 3.5 years of follow-up is not the same for both sexes: for males it is 5.3%, while for females it is 7.4%. [3,4] Applying the same assumptions, we would now "only" need to include 5528 female patients in our trial; however, based on the male event rate, 7870 males should be included (Table 1). Thus, if the

sample size (N=6282) calculated on the prevalence in the general population had been used, enough patients would have been included based on the female incidence of the endpoint, but with the male incidence of the endpoint the study would be underpowered.
The risk of all-cause death is high in males (though not significantly), and this may lead to a competing risk of death which prevents male AF patients developing a stroke. [3] When assessing the effect of a new treatment in these patients, a competing-risk assessment is needed. In addition, the higher risk of all-cause death in males with AF means that fewer males will complete the follow-up, so less data will be collected for male patients. Another point is the timing for the outcome event to present itself. Figure 1 shows the risk for stroke in AF patients over time by sex. [4] Females reach a cumulative risk of 10% in little over six years while males take more than ten years. These differences in time-to-event have to be taken in consideration when designing trials which include both males and females.

FIGURE 1 - Kaplan-Meier curves for stroke in AF patients, adapted from Marzona et al., 2020.



In conclusion, when assessing the effect of a treatment in a population including both sexes, a researcher should be aware of what proportion of the population may have competing risks. In addition, the timing of outcome events during follow-up should be borne in mind.

References

1. *Feinberg WM, Blackshear JL, Laupacis A, Kronmal R, Hart RG. Prevalence, age distribution, and gender of patients with atrial fibrillation. Analysis and implications. Arch Intern Med. 1995;155:469–73.*

2. Wagstaff AJ, Overvad TF, Lip GYH, Lane DA. Is female sex a risk factor for stroke and thromboembolism in patients with atrial fibrillation? A systematic review and meta-analysis. QJM. 2014;107:955–67.

3. Marzona I, Proietti M, Farcomeni A, Romiti GF, Romanazzi I, Raparelli V, et al. Sex differences in stroke and major adverse clinical events in patients with atrial fibrillation: A systematic review and meta-analysis of 993,600 patients. International Journal of Cardiology. 2018;269:182–91.

4. Marzona I, Proietti M, Vannini T, Tettamanti M, Nobili A, Medaglia M, et al. Sex-related differences in prevalence, treatment and outcomes in patients with atrial fibrillation. Intern Emerg Med. 2020;15:231–40.

Acknowledgements

The authors would like to thank Judith Baggot for the revision of the text.

