Journal of Molecular Medicine (2019) 97:871–877
https://doi.org/10.1007/s00109-019-01774-0

JMolMed

ORIGINAL ARTICLE

# Benefits of a factorial design focusing on inclusion of female and male animals in one experiment

Thorsten Buch[1] · Katharina Moos[2,3] · Filipa M. Ferreira[1] · Holger Fröhlich[4] · Catherine Gebhard[5] · Achim Tresch[2,3]

Received: 1 December 2018 / Revised: 5 February 2019 / Accepted: 10 March 2019 / Published online: 13 April 2019


## Abstract

Disease occurrence, clinical manifestations, and outcomes differ between men and women. Yet, women and men are most of the time treated similarly, which is often based on experimental data over-representing one sex. Accounting for persisting sex bias in biomedical research is the misconception that the analysis of sex-specific effects would double sample size and costs. We designed an analysis to test the potential benefits of a factorial study design in the context of a study including male and female animals. We chose a 2 × 2 factorial design approach to study the effect of treatment, sex, and an interaction term of treatment and sex in a hypothetical situation. We calculated the sample sizes required to detect an effect of a given magnitude with sufficient power and under different experimental setups. We demonstrated that the inclusion of both sexes in experimental setups, without testing for sex effects, requires no or few additional animals in our scenarios. These experimental designs still allow for the exploration of sex effects at low cost. In a confirmatory instead of an exploratory design, we observed an increase in total sample sizes by 33%, at most. Since the complexities associated with this mathematical model require statistical expertise, we generated and provide a sample size calculator for planning factorial design experiments. For the inclusion of sex, a factorial design is advisable, and a sex-specific analysis can be performed without excessive additional effort. Our easy-to-use calculation tool provides help in designing studies with both sexes and addresses the current sex bias in preclinical studies.

## Key messages

- Both sexes should be included into animal studies.
- Exploratory study of sex effects can be conducted with no or small increase in animal number.
- Confirmatory analysis of sex effects requires maximum 33% more animals per study.
- Our calculation tool supports the design of studies with both sexes.

**Keywords** Sex · Animal experimentation · Factorial design · Power

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s00109-019-01774-0) contains supplementary material, which is available to authorized users.

✉ Thorsten Buch
thorsten.buch@uzh.ch

✉ Achim Tresch
achim.tresch@uni-koeln.de

[1] Institute of Laboratory Animal Science, University of Zurich, Wagistrasse 12, 8952 Schlieren, Zurich, Switzerland

[2] Institute of Medical Statistics and Computational Biology, Faculty of Medicine, University of Cologne, Bachemer Str. 86, 50931 Cologne, Germany

[3] Center for Data and Simulation Science (CDS), University of Cologne, Cologne, Germany

[4] Bonn-Aachen International Center for IT (b-it), University of Bonn, Bonn, Germany

[5] Center for Molecular Cardiology, University of Zurich, Zurich, Switzerland

## Introduction

Differences between the sexes have long been neglected in biomedical research. This was often based on the assumption that results from one sex can simply be applied to the other sex. Concerns that the estrous cycle in females might increase phenotypic variance and thereby decrease the statistical power, which is needed to detect the primary effect, then led to an over-reliance on male subjects [1]. Both assumptions have essentially been disproven [2–4]. Indeed, an increasing number of studies indicate that sex differences are substantial in drug responses, and also in cardiovascular, neurological, and autoimmune disorders [5, 6]. These studies have raised awareness that the over-reliance on either male or female research subjects obscures key sex differences. The resulting lack of data hampers the accurate design of subsequent clinical trials. Therefore, major funding agencies addressed this bias and called for the analysis of any sex difference. Accordingly, in 2014, the US National Institutes of Health (NIH) developed the sex as a biological variable (SABV) policy requiring sex- or gender-specific reporting of research and issued a mandate for including both sexes in all vertebrate studies [7, 8]. The Canadian Institutes of Health Research [9], the European Commission's Horizon 2020 research program [10], and the gender policy committee of the European Association of Science Editors made similar announcements [11, 12]. Despite these policies, many researchers in various disciplines still conduct experiments with only one sex [13–19]. One prominent reason for the use of only one sex in basic and translational animal research is the pressure to reduce the number of animals according to the "3R" principle (3R = reduction, refinement, replacement) [20]. Thus, scientists currently find themselves in the dilemma of having to minimize animal consumption while still detecting a treatment effect and having to maximize insight by including both sexes.

A study setup based on a conventional randomized single-factor design does indeed require the duplication of sample size in order to assess sex-specific effects. In contrast, the so-called (*balanced*) *factorial design*, which uses the analysis of variance (ANOVA), offers the possibility of analyzing the influence of more than one categorical variable on the study outcome. Previous work suggests that including an additional experimental variable in a factorial design may yield supplemental information without causing a major increase in the number of animals, thereby serving the 3R principle [21–24]. Taking on the current challenge of implementing sex-sensitive findings into research strategies, we investigated the benefits of a factorial design in the context of a hypothetical study with male and female animals. We assessed the effect of sex as a second factor variable on the statistical power and the number of animals required.

## Results

### The inclusion of sex as a second factor requires no or only a few additional animals

For clarity in presentation, we restricted ourselves to the situation in which the primary objective is the identification of a treatment effect. One detects such an effect using a 1-factor (1f) ANOVA when comparing the observed outcomes of a group of animals that received treatment with the outcomes of a control group (either without treatment, with placebo, or alternative drug treatment). The question we wanted to answer is if it is recommendable to introduce a second factor, sex, to the design of an animal experiment. Therefore, we discuss a $2 \times 2$ factorial design, in which we tested the influence of two binary factor variables, treatment and sex, on a quantitative outcome by use of 2-factor (2f) ANOVA. We will briefly touch on possible extensions to more than 2 factors in the "Discussion."

ANOVA quantifies three effects: two *main effects* describing the average effect of treatment and the average effect of sex, and the *interaction effect* describing the difference of males and females in their response to treatment. When we are only interested in testing for drug efficacy, we would perform only one test for the treatment effect (1f-ANOVA). By performing a 2f-ANOVA, we gained additional information about the existence and magnitude of the other two effects. Of note, $p$ values were not corrected for multiple testing and may only serve for data exploration. We calculated the number of animals needed in a $2 \times 2$ factorial design and compared them to a 1f-ANOVA design (Fig. 1). We did this for various scenarios, using relative effect sizes of 0.4, 0.5, and 1, a test power in the range of 80–90%, and a fixed significance level of 5% (see Box I for a definition of these quantities). The inclusion of sex as an additional factor did not lead to an increase in the total sample sizes required for large relative effect sizes and moderate power, compared to the 1f-ANOVA. For smaller effect sizes or higher power, the increase was small (at most 2 animals) (Table 1). Thus, restricting the analysis to the primary factor, the inclusion of sex led to no or only a few additional animals. This type of analysis allows for an exploratory post hoc determination of a possible sex effect and provides a basic concept for further studies.

### Analysis of one additional factor, such as sex, leads to a minimum of extra cost

When testing whether not only the primary factor but also sex influences the outcome parameter, the analysis of an additional factor requires the application of a multiple testing correction (e.g., Bonferroni correction). This correction influences the sample sizes and the power of the tests. Figure 2 shows how total sample sizes change in the

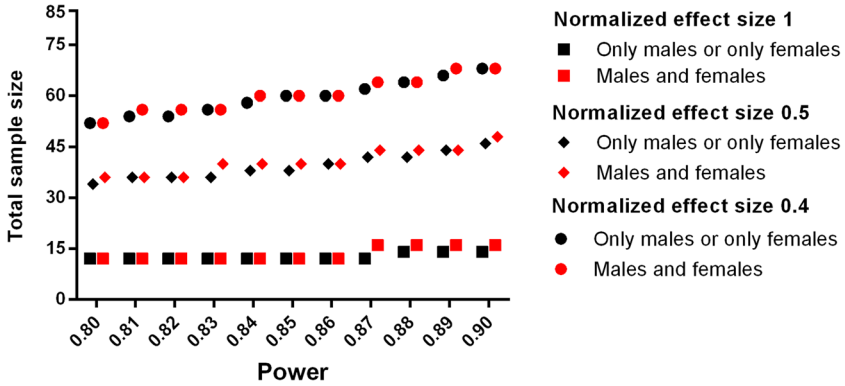


**Fig. 1** Including sex as confounder in absence of interaction leads to insignificant increases in total sample sizes. Comparison of total sample sizes between analysis without inclusion of sex or with inclusion of sex as a confounder. The modeled data has no interactions between the factors studied; the ANOVA model assumes no interactions. The data is solely analyzed for the primary factor, and no correction for multiple testing was performed (sex not analyzed). For each effect size, 1 or 2 factors (e.g., primary factor plus/minus sex) were modeled

scenarios described in the previous paragraph. The inclusion of a second factor and analysis by 2f-ANOVA with a Bonferroni correction did not result in a substantial increase in the number of animals required. The increase ranged from 2 to 16 animals, or 14 to 33%, as compared to 1f-ANOVA (Table 1). This type of analysis of sex effects, however, bears a considerable problem: any potential interaction between sex and the primary factor is being ignored. Instead one should use a model that estimates the interaction between sex and the investigated parameter as described in the following paragraph.

## The inclusion of interaction effects is mandatory

The analysis of both sexes can be considered as good scientific practice, even under consideration that only the minimum number of animals should be used. In this form of analysis, an interaction between sex and the primary factor can usually not be excluded. Such interaction is commonly known as synergism and means that the combined effect of the two factors is greater (or smaller) than the plain sum of their individual effects. Awareness of sex-dependent synergisms is important, as in the worst case a drug may be considered safe in one sex

**Table 1** Sample size calculations in different experimental setups

| Experiment | Factor | | Relative effect size | Total sample size | | | Sample size increase compared to 1f-ANOVA [%] | | |
|---|---|---|---|---|---|---|---|---|---|
| | Treatment | Sex | | Power 0.8 | Power 0.85 | Power 0.9 | Power 0.8 | Power 0.85 | Power 0.9 |
| 1f-ANOVA | C | – | 1 | 12 | 12 | 14 | – | – | – |
| | | | 0.5 | 34 | 38 | 46 | – | – | – |
| | | | 0.4 | 52 | 60 | 68 | – | – | – |
| 2f-ANOVA without interaction | C | e | 1 | 12 | 12 | 16 | 0 | 0 | 14 |
| | | | 0.5 | 36 | 40 | 48 | 6 | 5 | 4 |
| | | | 0.4 | 52 | 60 | 68 | 0 | 0 | 0 |
| 2f-ANOVA without interaction | C | c | 1 | 16 | 16 | 16 | 33 | 33 | 14 |
| | | | 0.5 | 44 | 48 | 56 | 29 | 26 | 22 |
| | | | 0.4 | 64 | 72 | 84 | 23 | 20 | 24 |
| 2f-ANOVA with interaction | C | e | 1 | 12 | 12 | 16 | 0 | 0 | 14 |
| | | | 0.5 | 36 | 40 | 48 | 6 | 5 | 4 |
| | | | 0.4 | 52 | 60 | 68 | 0 | 0 | 0 |
| 2f-ANOVA with interaction | C | c | 1 | 16 | 16 | 16 | 33 | 33 | 14 |
| | | | 0.5 | 44 | 48 | 56 | 29 | 26 | 22 |
| | | | 0.4 | 64 | 72 | 84 | 23 | 20 | 24 |

Comparison of sample sizes for different experimental setups. Factors were either tested confirmatory (c) or exploratory (e). Bonferroni correction is always applied if both factors are tested confirmatory. Setups using ANOVA without interactions assume data without interactions as well

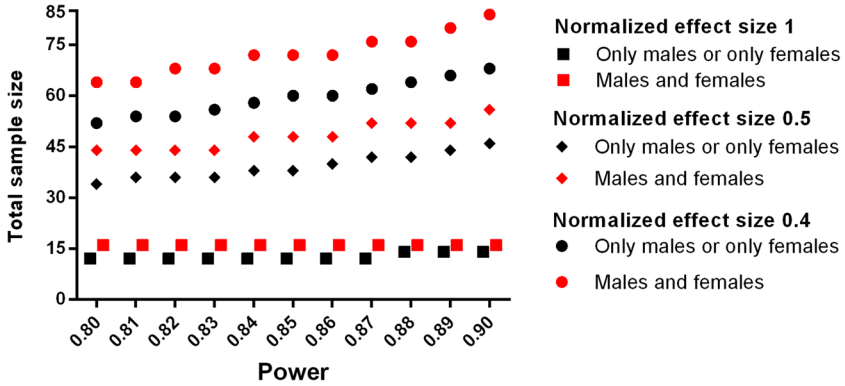


**Fig. 2** Inclusion of sex as analyzed variable in absence of interaction leads to moderate increases in total sample sizes. Comparison of total sample sizes between analysis without inclusion of sex or with inclusion of sex as a second analyzed variable without interaction effects. The modeled data has no interactions between the factor studied and sex. The ANOVA model assumes no interactions between the factors. A post hoc correction (Bonferroni) for multiple comparison was performed to include sex as analyzed variable. For each effect size, 1 or 2 factors were modeled

while causing detrimental side effects in the opposite sex. Therefore, designs of good animal experiments need to reveal these effects.

We compared the power of 2f-ANOVA without interaction and 2f-ANOVA with interaction, applied to data with interaction effects. In this scenario, we found a noticeable loss of power in the 2f-ANOVA without interaction (Fig. 3a). This loss depended on the size of the interaction effect, with higher power loss correlating with stronger interactions. The reason is that a model which lacks interactions incorrectly interprets differences between groups, due to interactions, as noise. The only way to overcome this loss of power is to include interactions in the model. Next, we modeled the total sample sizes needed for 2f-ANOVA with and without interaction applied on data that show no interaction between treatment and sex. The 2f-ANOVA with interaction required a larger total sample

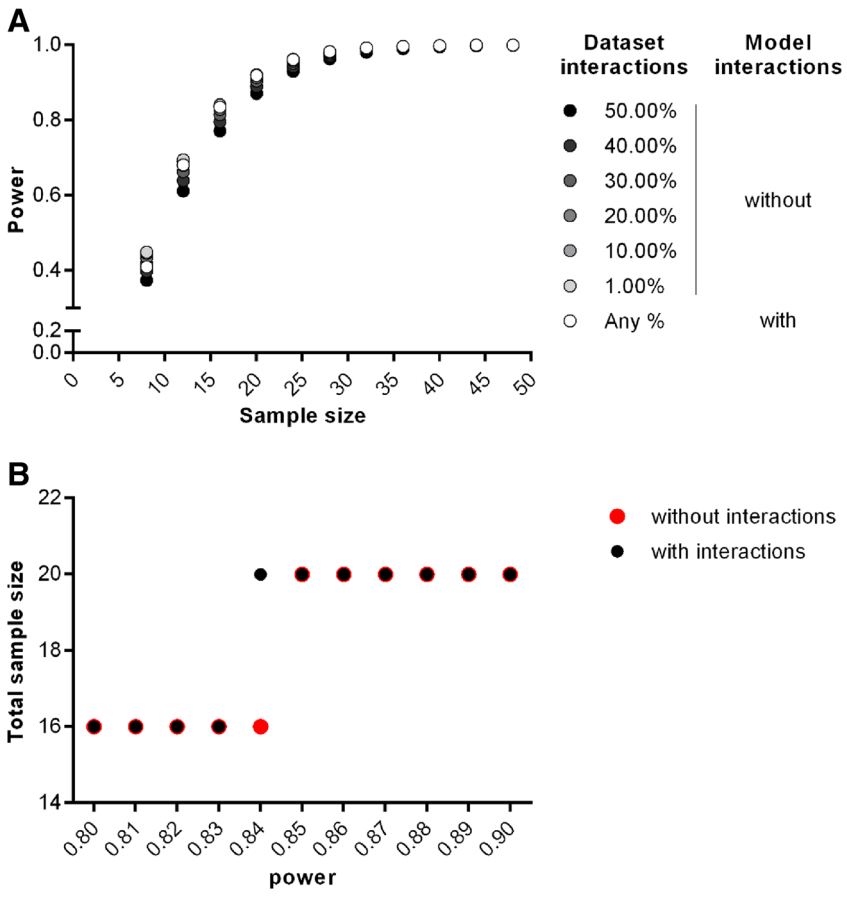


**Fig. 3** Exclusion of interaction in the ANOVA model decreases power, but its inclusion comes with little cost. 2f-ANOVA with or without interaction was used for the modeling. The interaction effect of the data set varies from 1% (light gray) to 50% (black). The second factor (sex) was treated as confounder. **a** Power versus total sample size, based on the data set with interactions. ANOVA with interaction (white dots) and without interaction (light gray to black dots). **b** Total sample size needed in dependence of power, based on a data set *without* interactions using a model with interaction. *ANOVA with interaction* (black dots) or no interaction effect (red dots, correct situation) were modeled

size only for an intermediate power of 0.84 (Fig. 3b). In comparison to a 1f-ANOVA of an experiment with one sex only, a 2f-ANOVA with interaction on a two-sex experiment, sex as confounder, requires at most two additional animals, amounting to a maximum increase of 14% in the total number of animals (Table 1).

Taken together, applying 2f-ANOVA without interaction in a situation where interactions cannot safely be excluded results in a major drop of power. This means that one might fail to find true effects of the primary factor or sex. On the other hand, applying 2f-ANOVA with interaction on data where no interactions are present comes along with little to no additional costs. Thus, the interaction model is clearly preferable.

## Testing for sex effects considering sex-specific interactions comes at reasonable costs

Next, we calculated the number of animals required for an experiment investigating the primary factor and sex effects in an appropriately powered confirmatory experiment using 2f-ANOVA with interaction and Bonferroni correction. We compared this setup with an experimental one in which only one sex, either male or female, was included and analyzed by use of 1f-ANOVA. Again, we investigated three different effect sizes (remark: the size of the interaction effect does not influence the total sample size in the 2f-ANOVA with interaction). In our simulations with fixed effect sizes, significance level, and power, the increase in total sample size amounted to 2–16 animals, or 14–33% (Fig. 4 and Table 1).

## Recommendations for the researcher

According to current policies, the inclusion of both sexes should become the standard procedure for all experimental studies. Therefore, we recommend using one of the following approaches:

- To perform a sex-aware analysis of the primary factor effect, use 2f-ANOVA with interaction but without Bonferroni correction. Estimated sex effects are of exploratory nature and may guide further studies.
- To perform a confirmatory analysis of both the primary factor and sex effect, use 2f-ANOVA with interactions and with Bonferroni correction.

Bear in mind that it is mandatory to choose and fix the analysis approach before having seen the data. Trying several approaches and choosing "the best" afterwards will lead to a violation of the significance level of the findings (over-optimism, "p-hacking").

One should further keep in mind the following general points:

- The application of ANOVA requires the data in each group to follow approximately normal distribution with identical variance. Verifying (e.g., using historical data) that an assumption is reasonable for the experiment in question is vital.
- Since within-group variance, a crucial parameter of the power calculation, is notoriously hard to guess, one should use a conservative estimate (i.e., a number that rather overestimates the within-group variance).
- If sample sizes rule out an analysis in a single batch, it is mandatory to use an experimental design that avoids biases due to batch effects. For example, if you have a sample size of 15 and perform the analysis on 3 days in batches/blocks of 5, make sure to distribute each factor (treatment/control, male/female) as evenly as possible across all batches [25].

## Power calculations for 2f-ANOVA made easy

Several tools provide researchers with support for power and sample size calculations [26–30]. However, many of these tools are restricted to 1f-ANOVA or are difficult to fill-in and interpret by the practitioners.

To allow for easy power calculations for a 2f-ANOVA, we implemented an Excel sheet for two factors

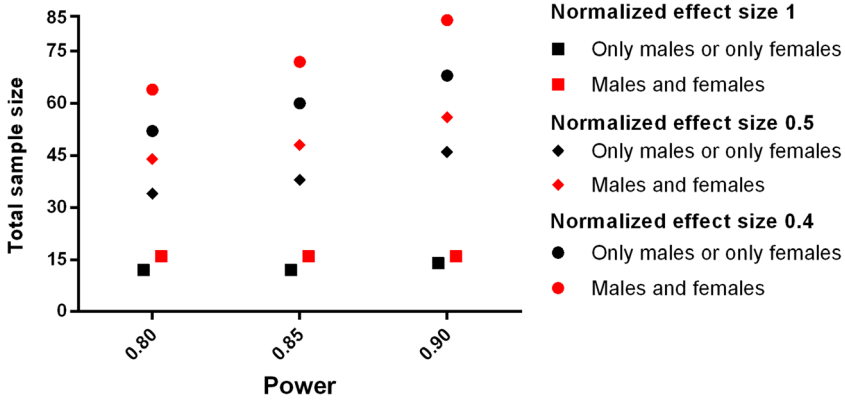


**Fig. 4** Analyzing sex effects including interaction moderately increases sample sizes. Comparison of sample sizes for ANOVA with interaction and with Bonferroni correction to sample sizes of a one-sex analysis. Total sample size needed in dependence of power. One-sex analysis (black dots) versus two-sex analysis (red dots)

(Supplemental Material SX). It depends on the Excel Package "Real Statistics" [31] which can be installed using the instructions on the package's webpage. The experimenter has to enter the expected mean value for the observations in each experimental group, the expected variance within the groups, as well as the desired significance and power level. Additionally, the user has to specify which effects are to be tested confirmatory and which should be considered exploratory. The total sample sizes are then computed according to the user's input (multiple testing correction is automatically applied, if necessary). We report the dependence of required total sample sizes as a function of variance and power in additional table sheets.

Box I

| | |
|---|---|
| ANOVA | Analysis of variance; a statistical test to check for the presence of differences within multiple levels of a factor (ANOVA does not analyze between which or how many levels the differences occur); if 1 (2, 3, …) factors are tested, it is also referred to as 1 (2, 3, …)—factorial ANOVA |
| Factor | A variable being analyzed (e.g., age, sex, type of treatment) |
| Factor level | Manifestation of a variable (e.g., sex factor with factor levels, male and female) |
| Relative effect size (often also termed signal-to-noise ratio) | The absolute difference of the group mean (here, the mean of all samples with the same sex and same treatment) from the overall mean, normalized by within-group standard deviation. |
| Power | True positive rate; the probability that a difference in the data is detected by a statistical test (e.g., in 10 cases with differences in the data, a test with a power of 0.8 (or 80%) detects 8 cases as statistically significant) |
| Significance level | False positive rate; the probability that, when there is no difference in the data, a statistically significant difference is called by the test (e.g., in 100 cases with no differences in the data, a test with a significance level of 0.05 (or 5%) calls 5 cases as statistically significant) |
| Multiple testing | Data is analyzed using multiple statistical tests; the overall significance level of all tests will be higher than that of the single tests. This is usually corrected (e.g., when performing two tests, both with a 5% significance level, the overall significance level is $1-(0.95\times0.95)=0.0975$ which is 9.75%) |

For details on the used ANOVA formulas and the scripts used for this study, please refer to Supplementary Methods and Supplementary Scripts, respectively.

## Discussion

Integrating sex as a factor into biomedical research is an essential step towards precision medicine and tailored care for women and men. The use of a factorial design in animal experiments has been brought forward as an appropriate method to include sex-specific information in experimental setups (24). Indeed, in a retrospective investigation of published experimental data, Festing showed that all the experiments he assessed had flawed statistical analyses and would have benefited from the application of a factorial design and ANOVA (15).

While such reanalysis of experimental data draws attention to the problem of an inappropriate use of statistical tools and (over)use of animals, it does not inform about the applicability of a 2f-ANOVA design. We therefore systematically assessed the advantages and disadvantages of a factorial design and the use of ANOVA for the inclusion of sex as an additional factor through in silico modeling. We compared one-factorial designs with designs including sex as second factor. With one-factorial design, testing different sexes means one would have to perform independent experiments for each sex, therefore doubling the number of animals used, without providing information about interaction effects. Contrary to Festing's findings (25), the addition of an additional factor (sex) is not without any cost: it yields a very modest increase in total sample sizes. This increase was smaller when sex was only investigated in an exploratory rather than a confirmatory manner since a multiple testing correction needs to be applied in the latter case. Including interaction in the 2f-ANOVA model is of particular importance as even the presence of small interactions may influence the effect estimates and may lead to the misinterpretation of data. Considering the marginal increases in sample sizes and thus animal consumption when an additional factor is included in an experiment, we recommend the inclusion of sex as a second factor.

Of note, similar considerations hold for ANOVA designs with more than 2 factors. However, these designs are less common, since they are more difficult to implement in practice: Typically, not all animals can be analyzed in one batch. As pointed out in the "Recommendations for the researcher," the assignment of samples to batches is a non-trivial task, since one has to avoid confounding batch effects. Nevertheless, ANOVA designs with more than 2 factors can have their merits in well-planned experiments.

To the ease of the experimentalist, we provide an Excel sheet for power calculations in 2f-ANOVA designs in the Supplementary Material. Thus, we offer a strategy to

successfully design studies including both sexes while conforming to the requirements of ethical principles and financial limitations. Our approach may foster initiatives that incorporate sex in experimental studies with the ultimate goal to improve trial design and develop optimal treatment strategies for both women and men.

**Author contribution** Conceptualization and supervision were performed by TB and AT. Methodology and formal analysis were performed KM. Verification was performed by HF. Visualization was performed by FMF and KM. The original draft was written by TB, FMF, CG, KM, and AT. The manuscript was reviewed and edited by TB, FMF, HF, CG, KM, and AT.

**Funding** TB was supported by the Hertie Foundation and the Swiss Multiple Sclerosis Society. CG was supported by grants from the Swiss National Science Foundation, the Olga Mayenfisch Foundation, Switzerland, the OPO Foundation, Switzerland, the Novartis Foundation, Switzerland, the Swissheart Foundation, and the Helmut Horten Foundation, Switzerland. KM was supported by the German Research Foundation (DFG).

## References

1. Wizeman TM, Pardue ML (2001) Exploring the biological contributions to human health: does sex matter? In: National Academy Press. USA, Washington, DC
2. Itoh Y, Arnold AP (2015) Are females more variable than males in gene expression? Meta-analysis of microarray datasets. Biol Sex Differ 6:18
3. Dayton A, Exner EC, Bukowy JD, Stodola TJ, Kurth T, Skelton M, Greene AS, Cowley AW Jr (2016) Breaking the cycle: estrous variation does not require increased sample size in the study of female rats. Hypertension (Dallas, Tex : 1979) 68(5):1139–1144
4. Beery AK (2018) Inclusion of females does not increase variability in rodent research studies. Curr Opin Behav Sci 23:143–149
5. Regitz-Zagrosek V (2014) Sex and gender differences in pharmacotherapy. Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz 57(9):1067–1073
6. Beery AK, Zucker I (2011) Sex bias in neuroscience and biomedical research. Neurosci Biobehav Rev 35(3):565–572
7. New supplemental awards apply sex and gender lens to NIH-funded research. (2014). https://www.nih.gov/news-events/news-releases/new-supplemental-awards-apply-sex-gender-lens-nih-funded-research. Accessed 10.06.2018 2018
8. Clayton JA, Collins FS (2014) Policy: NIH to balance sex in cell and animal studies. Nature 509(7500):282–283
9. Research CIoH (2018) How to integrate sex and gender into research. http://www.cihr-irsc.gc.ca/e/50836.html
10. Guidance on Gender Equality in Horizon 2020 (2016). Version 2 edn.,
11. Committee GP (2018) Gender Policy Committee. The European Association of Science Editors (EASE). http://www.ease.org.uk/strategy-groups/gender-policy-committee/. Accessed 11.06.2018 2018
12. Heidari S, Babor TF, De Castro P, Tort S, Curno M (2016) Sex and gender equity in research: rationale for the SAGER guidelines and recommended use. Res Integr Peer Rev 1:2
13. Bryant J, Yi P, Miller L, Peek K, Lee D (2018) Potential sex Bias exists in orthopaedic basic science and translational research. J Bone Joint Surg Am 100(2):124–130
14. Florez-Vargas O, Brass A, Karystianis G, Bramhall M, Stevens R, Cruickshank S, Nenadic G (2016) Bias in the reporting of sex and age in biomedical research on mouse models. eLife 5. https://doi.org/10.7554/eLife.13615
15. Potluri T, Engle K, Fink AL, Vom Steeg LG, Klein SL (2017) Sex reporting in preclinical microbiological and immunological research. mBio 8(6). https://doi.org/10.1128/mBio.01868-17
16. Ramirez FD, Motazedian P, Jung RG, Di Santo P, MacDonald ZD, Moreland R, Simard T, Clancy AA, Russo JJ, Welch VA, Wells GA, Hibbert B (2017) Methodological rigor in preclinical cardiovascular studies: targets to enhance reproducibility and promote research translation. Circ Res 120(12):1916–1926
17. Stephenson ED, Farzal Z, Kilpatrick LA, Senior BA, Zanation AM (2018) Sex bias in basic science and translational otolaryngology research. Laryngoscope. https://doi.org/10.1002/lary.27498
18. Will TR, Proano SB, Thomas AM, Kunz LM, Thompson KC, Ginnari LA, Jones CH, Lucas SC, Reavis EM, Dorris DM, Meitzen J (2017) Problems and progress regarding sex Bias and omission in neuroscience research. eNeuro 4(6). https://doi.org/10.1523/eneuro.0278-17.2017
19. Yoon DY, Mansukhani NA, Stubbs VC, Helenowski IB, Woodruff TK, Kibbe MR (2014) Sex bias exists in basic science and translational surgical research. Surgery 156(3):508–516
20. Russell WMS, Burch RL, Hume CW (1959) The principles of humane experimental technique
21. Fisher RA (1935) The design of experiments. Oliver and Boyd
22. Festing MF (1994) Reduction of animal use: experimental design and quality of experiments. Lab Anim 28(3):212–221
23. Festing MF (1992) The scope for improving the design of laboratory animal experiments. Lab Anim 26(4):256–268
24. Miller LR, Marks C, Becker JB, Hurn PD, Chen WJ, Woodruff T, McCarthy MM, Sohrabji F, Schiebinger L, Wetherington CL, Makris S, Arnold AP, Einstein G, Miller VM, Sandberg K, Maier S, Cornelison TL, Clayton JA (2017) Considering sex as a biological variable in preclinical research. FASEB J 31(1):29–34
25. Montgomery DC (2012) Design and analysis of experiments. Wiley, Hoboken
26. Faul F, Erdfelder E, Buchner A, Lang AG (2009) Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 41(4):1149–1160
27. Champely S, Ekstrom C, Dalgaard P, Gill J, Weibelzahl S, Anandkumar A, Ford C, Volcic R, HDe Rosario H (2018) Basic functions for power analysis. https://github.com/heliosdrm/pwr. Accessed 11.06.2018
28. Fan FY (2017) Basic functions for power analysis and effect size calculation. https://cran.r-project.org/web/packages/powerAnalysis/index.html. Accessed 11.06.2018 2018
29. StatTools : Resource Index (Subjects). (2014) Chinese University of Hongkong: Department of Obstretics and Gynaecology. http://www.obg.cuhk.edu.hk/ResearchSupport/StatTools/ResourceIndex_Subjects.php. Accessed 11.06.2018 2018
30. Bioinformatics Q (2018) Power or sample size calculator. https://www.anzmtg.org/stats/PowerCalculator. Accessed 11.06.2018 2018
31. Zaiontz C (2018) Real statistics using Excel. www.real-statistics.com. Accessed 11.06.2018 2018