Tidsskriftet Tidsskrift for Den norske legeforening

Percentages in rows or columns?

MEDICINE AND NUMBERS

STIAN LYDERSEN

stian.lydersen@ntnu.no
Stian Lydersen, PhD, professor of medical statistics at the Regional Centre for Child and Youth Mental Health and Child Welfare (Central Norway), Department of Mental Health, Norwegian University of Science and Technology.
The author has completed the ICMJE form and declares no conflicts of interest

When reporting counts and comparing proportions, it can be useful to present the figures in a contingency table. But should the percentages be reported in the rows or the columns?

To answer this question, we can ask ourselves what is *exposure* and what is *outcome*, and report percentages within each exposure group. The normal approach is to show exposure in the rows, and thereby also percentages in the rows (1, p. 85).

Exposure or outcome?

Let us consider an example. Gårdvik et al. report results from a follow-up study of individuals three years after their period as outpatients in a mental health clinic for children and youth (2). Of 464 persons, 256 were women, 109 of whom were diagnosed with anxiety at the time of follow-up, while 147 were not. Of 208 men, a total of 43 were diagnosed with anxiety, while 165 were not (Table 2 in (2)). These figures can be cross-tabulated as shown in Table 1. Here, the percentages of persons with and without a diagnosis within each gender are shown first, i.e. percentages in the rows of the tabular presentation. In this study, 43 % of the women and 21 % of the men were diagnosed with anxiety.

Table 1

Contingency table of persons diagnosed and not diagnosed with anxiety, by gender (from Table 2 in (2)). Percentages within each gender are placed to the right of the counts, while percentages within each diagnosis group are placed below the counts.

	Diagnosed with anxiety	No diagnosis	Total
Women	109 (43 %) (72 %)	147 (57 %) (47 %)	256 (100 %)
Men	43 (21 %) (28 %)	165 (79 %) (53 %)	208 (100 %)
Total	152 (100 %)	312 (100 %)	464

Could it alternatively be relevant to show percentages within each column? These percentages are also shown in Table 1: Among those who had the diagnosis, 72 % were women, and among those who did not, 47 % were women.

The two alternatives elucidate two different questions: a) What proportion of the men and women in the study have the diagnosis?; and b) What are the proportions of men and women in the groups with and without the diagnosis?

In this case, it is most natural to think of gender as exposure and the diagnosis of anxiety as outcome, rather than the opposite, and a) will be the natural question to ask. It will thus be appropriate to report percentages within each gender, i.e. in the rows, as shown in the table. Table 1 also includes 100 % for the total, to make it clear to the reader whether the percentages are reported in the rows or the columns.

Background data in randomised trials

When showing a table of background data in a randomised, controlled trial, it might be more practical to show the *exposure*, that is, *the treatment group*, in the columns. Table 2 shows an excerpt from a table of descriptive background data for a trial involving exercise for participants aged 70 to 75 years (3). The proportions are calculated within each intervention group (column), and are almost the same across the groups, as can be expected in a randomised trial. In Table 2, rows of redundant information have been omitted, for example that the number of men in the control group was 780–401 = 379 (48.6 %).

Table 2

Background data in a randomised, controlled trial of exercise for participants aged 70 to 75 years. Excerpt from Table 1 in (3)).

Background variable	Control *n* = 780	Moderate intensity exercise *n* = 387	High intensity exercise *n* = 400
Women	401 (51.4 %)	199 (51.4 %)	190 (47.5 %)

Background variable	Control n = 780	Moderate intensity exercise n = 387	High intensity exercise n = 400
Cardiovascular disease	136 (17.4 %)	77 (19.9 %)	60 (15.0 %)

When presenting a contingency table, the choice of whether to report percentages in the rows or the columns should be made deliberately. The choice should reflect what is considered to be *exposure*, as well as what is useful information for the reader. Unfortunately, we regularly see presentations where the authors have not been conscious of this.

REFERENCES

1. Fagerland M, Lydersen S, Laake P. Statistical Analysis of Contingency Tables. Boca Raton, FL: CRC Press, 2017.

2. Gårdvik KS, Rygg M, Torgersen T et al. Psychiatric morbidity, somatic comorbidity and substance use in an adolescent psychiatric population at 3-year follow-up. Eur Child Adolesc Psychiatry 2021; 30: 1095–112. [PubMed] [CrossRef]

3. Stensvold D, Viken H, Steinshamn SL et al. Effect of exercise training for five years on all cause mortality in older adults-the Generation 100 study: randomised controlled trial. BMJ 2020; 371: m3485. [PubMed][CrossRef]

Publisert: 9 September 2024. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.24.0329