BMJ 2014;349:g7327 doi: 10.1136/bmj.g7327 (Published 28 November 2014)

ENDGAMES

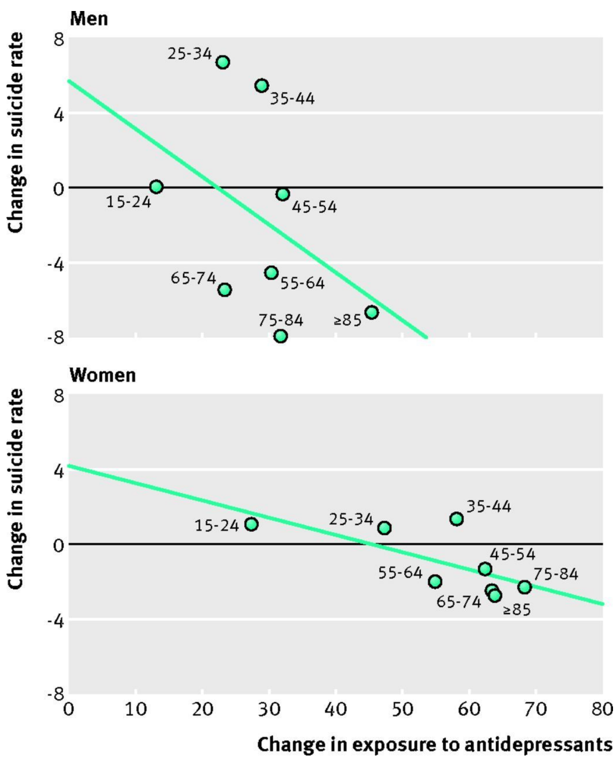
Spearman's rank correlation coefficient

Philip Sedgwick *reader in medical statistics and medical education*

Institute for Medical and Biomedical Education, St. George's, University of London, UK

Researchers investigated the association between suicide rates and antidepressant prescribing in Australia.[1] A retrospective analysis of national databases between 1991 and 2000 was undertaken. Participants were aged 15 years or more. Rates of suicide and antidepressant prescribing were recorded by sex and 10 year age groups (15-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84, and ≥85 years). Suicide rates were expressed as the number per 100 000 population for each of the five-year periods of 1986-90 and 1996-2000. The changes in suicide rates from 1986 to 90 to 1996-2000 were derived. Antidepressant prescribing was expressed as the estimated defined daily doses per 1000 population per day (DDD/1000/day). Defined daily dose was based on the assumed average daily dose of the drug when used by adults for its main indication. The change in antidepressant prescribing from 1991 to 2000 was derived.

The researchers reported that from 1986 to 90 to 1996-2000, suicide rates decreased in older men and older women, and increased in younger men and younger women. From 1991 to 2000, rates of antidepressant prescribing increased across all age groups in both men and women. The changes in suicide rates were plotted against those in antidepressant prescribing across age groups, with men and women investigated separately (fig 1). Spearman's rank correlation coefficient was used to measure the association between changes in suicide rates and antidepressant prescribing. For men and women, there was a negative correlation, with the largest declines in suicide in the age groups associated with the greatest increases in antidepressant prescribing. The association was significant in women (r_s=−0.74; P=0.04) but not in men (r_s=−0.62; P=0.10).



Scatter plot of change in suicide rates (100 000 population) (from 1986 to 90 to 1996-2000) against change in rates of antidepressant prescribing (DDD/1000/day) (from 1991 to 2000) for the 10 year age groups in men and women

The researchers suggested that an increase in antidepressant prescribing may be a proxy marker for improved overall management of depression. If so, increased prescribing of antidepressants in general practice may have a quantifiable benefit on the mental health of the population.

Which of the following statements, if any, are true?

a. Spearman's rank correlation coefficient provided a measure of the strength of a monotonic association between changes in suicide rates and antidepressant prescribing across the age groups

Correspondence to P Sedgwick p.sedgwick@sgul.ac.uk

b. The significance test for Spearman's rank correlation coefficient is parametric

c. It can be concluded that for women the decline across the age groups in suicide rates was caused by an increase in antidepressant prescribing.

Answers

Statement *a* is true, whereas *b* and *c* are false.

The association between changes in suicide rates and antidepressant prescribing across the age groups was represented in a scatter plot, with men and women plotted separately (fig 1). The researchers drew a straight line through the points on each scatter plot to suggest the most appropriate linear association.

Spearman's rank correlation coefficient measured the strength and direction of a monotonic association between changes in suicide rates and antidepressant prescribing across the age groups (*a* is true). A monotonic association is one where, as the value of one variable increases, so also does the value of the other, or as the value of one variable increases the other variable decreases. The variables will increase, or decrease, throughout the range of measured values but not necessarily at the same rate. An example of a monotonic association is the exponential curve. If an association was linear, then the variables would increase, or decrease, at the same rate throughout the range of measured values.

Since Spearman's rank correlation coefficient measures the strength and direction of a monotonic association, it was possibly inappropriate for the researchers to draw a straight line through the points on each scatter plot to indicate a linear association. Pearson's correlation coefficient, described in a previous question,[2] is used to measure the strength and direction of a linear association between two variables. In particular, Pearson's correlation coefficient measures how closely the points lay about the linear association in a scatter plot. Simple linear regression, described in previous questions,[3 4] would provide the gradient of a straight line drawn through the points in a scatter plot, and therefore a measure of the nature of the linear association.

Spearman's rank correlation coefficient, as for all correlation coefficients, is measured on a scale with no units with values ranging from −1 through 0 to +1. For the example above, the correlation coefficient was negative for both men and women, indicating that between 1991 and 2000 reductions in suicide rates were associated with increases in antidepressant prescribing across age groups. If a positive correlation had existed, then increases in suicide rates would have been associated with increases in antidepressant prescribing. The magnitude of the correlation coefficient indicates the strength and direction of the association; in general, it measures how closely the points lay about the monotonic association. A correlation coefficient closer to −1 or +1 indicates greater association, with values of +1 or −1 indicating an exact monotonic association. A correlation coefficient of zero would indicate that there was no association between the two variables—that is, they were not correlated.

Traditional statistical hypothesis testing with a null and alternative hypothesis[5] was undertaken, enabling a P-value to be derived to test the significance of the Spearman's rank correlation coefficients. Separate hypotheses tests were undertaken for men and women. The null hypothesis stated that the population correlation coefficient from which the sample was taken was zero. The alternative hypothesis stated that the population correlation coefficient from which the sample was taken was not equal to zero; the alternative hypothesis was two-sided, so the population correlation coefficient could be <0 or >0. A significant negative correlation was seen between the changes in suicide rates and antidepressant prescribing for women (P=0.04), but that for men did not reach significance (P=0.10).

The significance test for Spearman's rank correlation coefficient is non-parametric (*b* is false). In contrast, that for Pearson's correlation coefficient described above is parametric. Parametric and non-parametric tests have been described in a previous question.[6] Pearson's correlation coefficient is used when both variables are measured on a continuous scale, with the assumption that a linear association exists between them, plus at least one variable is distributed normally. In the example above, the significance test for Spearman's rank correlation coefficient made no distributional assumptions about either of the variables (changes in suicide rates or antidepressant prescribing). More generally, Spearman's rank correlation coefficient may be used if the assumptions for Pearson's correlation coefficient do not hold—that is, a linear association cannot be assumed—if neither variable is distributed normally, or if at least one variable is discrete (for example, the number of teeth extracted) or measured on an ordinal scale (for example, a depression rating score). The population parameter for Spearman's rank correlation coefficient is denoted by ρ_s and that for Pearson's correlation coefficient by ρ, whereas the sample estimates are represented by r_s and r, respectively.

The correlation between changes in suicide rates and antidepressant prescribing across age groups was significant for women. However, it cannot be inferred that for women an increase in antidepressant prescribing caused a decrease in suicide rates (*c* is false). Correlation investigates the association between two variables; it does not permit conclusions about causation to be inferred – that is, changes in one variable directly causes changes in the other. Furthermore, an ecological study design was used in the above study.[7] Because the unit of analysis was the age group—that is, the data were aggregated and analysed for each age group and not for each adult, the results of the study are prone to the ecological fallacy. The ecological fallacy is a term used when collected data are analysed at a group level and the results assumed to apply to associations at the individual level. In the above study, a negative correlation was reported between changes in suicide rates and antidepressant prescribing across age groups. Although those age groups with the greatest increase in antidepressant prescribing tended to show the greatest reduction in suicide rates, we cannot assume for any age group that an adult prescribed more antidepressants would be less likely to commit suicide.

The resulting P-value for a significance test for a correlation coefficient, whether parametric or non-parametric, is directly related to the sample size. In general, if the sample size is small, the correlation coefficient must be large (close to −1 or 1) for the association to be statistically significant. Conversely, if the sample size is large, the association may be statistically significant even if the value of the correlation coefficient is small and close to zero. There is a common misconception that a strong association exists between two variables only if the correlation coefficient is statistically significant. However, it is important to review the result of the significance test along with the value of the correlation coefficient, and to inspect the scatter plot of the two variables when investigating the strength of an association. In the study above, the correlation coefficient between changes in suicide rates and antidepressant prescribing were strong and of a similar magnitude for women (r_s=−0.74) and men (r_s=−0.62). However, the correlation for women was statistically significant (P=0.04) while that for men was not (P=0.10). The analyses for men and women were each based on a small sample size—that is, eight age groups. For that sample size a relatively large correlation coefficient was required to achieve statistical significance. These results therefore show why it may be misleading to make inferences about the strength of an association between two variables based on the statistical significance of the correlation coefficient.

Competing interests: none declared.

Provenance and peer review: not commissioned; externally peer reviewed.

1 Hall WD, Mant A, Mitchell PB, Rendle VA, Hickie IB, McManus P. Association between antidepressant prescribing and suicide in Australia, 1991-2000: trend analysis. *BMJ* 2003;326:1008. 10.1136/bmj.326.7397.1008 12742921
2 Sedgwick P. Pearson's correlation coefficient. *BMJ* 2012;345:e448310.1136/bmj.e4483.
3 Sedgwick P. Correlation versus linear regression. *BMJ* 2013;346:f268610.1136/bmj.f2686.
4 Sedgwick P. Simple linear regression. *BMJ* 2013;346:f234010.1136/bmj.f2340.
5 Sedgwick P. Understanding statistical hypothesis testing. *BMJ* 2014;348:g3557. 10.1136/bmj.g3557 24879817
6 Sedgwick P. Parametric v non-parametric statistical tests. *BMJ* 2012;344:e175310.1136/bmj.e1753.
7 Sedgwick P. Ecological studies: advantages and disadvantages. *BMJ* 2014;348:g2979. 10.1136/bmj.g2979 25134102