

Selected methods for the statistical analysis of qualitative data exemplified by the New York Heart Association and Canadian Cardiovascular Society classifications

Maria Sobol¹, Przemysław Kwasiborski²

¹Department of Biophysics and Human Physiology, Medical University of Warsaw

²Department of Biophysics and Human Physiology, Medical University of Warsaw, Department of Cardiology, Military Institute of Medicine, CSK MON, Warsaw

Kardiologia i Torakochirurgia Polska 2013; 10 (1): 84–86

Abstract

The aim of this paper is to revisit the statistical methods used in the analysis of qualitative data samples. Qualitative data are not measurable. The classifications used in cardiology, such as the Canadian Cardiovascular Society (CCS) classification or the New York Heart Association (NYHA) classes, constitute examples of such data. The CCS classification consists of four subclasses, to which patients are assigned on the basis of the subjective severity of angina pectoris. The NYHA classification consists of four subclasses as well, but uses the degree of physical activity limitation due to heart disease as its criterion. The choice of an appropriate method of statistical analysis depends, above all, on the correct assignment of the analyzed data to one of the two groups: qualitative or quantitative.

Key words: qualitative data, χ^2 test, CCS class.

Streszczenie

Niniejsza praca ma na celu krótkie przypomnienie i przeanalizowanie metod statystycznych służących do interpretacji wyników badań klinicznych charakteryzowanych za pomocą zmiennych jakościowych. Zmienne jakościowe są niemierzalne. Przykładem ich są coraz liczniejsze skale (klasy), wg których klasyfikuje się chorego kardiologicznego. Jest to np. klasa wg *Canadian Cardiovascular Society* (CCS) czy *New York Heart Association* (NYHA). Klasa CCS składa się z czterech podklas: CCS I, CCS II, CCS III i CCS IV, do których odpowiednio klasyfikuje się pacjenta ze względu na subiektywne nasilenie objawów dławicowych. Podobnie klasa NYHA składa się z NYHA I, NYHA II, NYHA III, NYHA IV, w której kryterium podziału jest subiektywne nasilenie dolegliwości związanych z niewydolnością serca. Wybór odpowiedniej metody analizy statystycznej zależy przede wszystkim od prawidłowego zakwalifikowania analizowanych danych do jednej z dwóch grup: ilościowych bądź jakościowych.

Słowa kluczowe: zmienne jakościowe, test χ^2 , klasa CCS.

The scales for classifying cardiac patients, which are becoming more and more widespread and numerous, may serve as an example of qualitative variables. Among the most commonly used classifications are the scales proposed by the New York Heart Association (NYHA), which use the subjective intensity of the ailments related to heart failure as its criterion [1], and the Canadian Cardiovascular Society (CCS), which takes into consideration the subjective severity of angina pectoris [2].

Additionally, when assessing patients after myocardial infarction, the Killip-Kimball [3] and Forrester [4] classifications are also used. Drawing conclusions concerning

relations between the groups of patients, established on the basis of the above classifications, requires the use of adequate statistical analysis methods, which often proves problematic for researchers. This is related to the fact that qualitative variables are assigned numerical values aside from descriptions. It needs to be taken into consideration that the numbers assigned to each class have no direct value reference (e.g. the four classes in the NYHA scale). Unfortunately, some researchers forget about this simple fact when conducting their statistical analyses. This is most likely the cause of a very common error found in numerous published works [5-8].

Address for correspondence: Maria Sobol, Warszawski Uniwersytet Medyczny, ul. Chałubińskiego 5, 02-004 Warszawa, tel. +48 22 628 63 34, Faks +48 22 628 78 46, e-mail: maria.sobol@wum.edu.pl

It is not correct to average the values of qualitative variables, as this leads to the use of inadequate statistical analysis and incorrect reasoning on its basis. Averaging is such a common mathematical operation that one does not consider the conditions that need to be met in order for the achieved results to be sensible and possible to interpret correctly.

Variables from different types of measurement scales are very often used in clinical research aimed at characterizing the studied patients. Among these, it is possible to distinguish the nominal scale (equal or different), the ordinal scale (higher or lower), the interval scale (how much higher), or the ratio scale (how many times higher).

The previously mentioned CCS scale is an ordinal scale. When using the ordinal scale, the researcher arranges the conducted observations (assigning ranks) with regard to a certain feature. These ranks refer to the position in a set of results arranged in ascending or descending order. Thus, in the case of the CCS classification, a group included in a study is assigned to one of the four subclasses, depending on the advancement of coronary heart disease. Individual classes are characterized by the following features: class I – angina only during strenuous physical activity; class II – slight angina during everyday activities, e.g. walking to the second floor or higher; class III - significant coronary ailments, e.g. when slowly walking up the stairs; class IV – angina during all physical activities or at rest. The NYHA scale (also known as the NYHA Functional Classification) is also an ordinal scale with four classes. The consecutive classes are characterized by the following clinical symptoms: NYHA class I – cardiac disease, but without any limitations in ordinary physical activity: everyday activities do not cause excessive fatigue, shortness of breath, angina, or palpitations; NYHA class II – cardiac disease with slight limitation during ordinary activity: excessive fatigue, angina, palpitations - not present at rest; NYHA class III – cardiac disease with marked limitation in everyday activity as a result of the ailments appearing even during slight physical activity, but not at rest; NYHA class IV – cardiac disease with even the slightest physical effort causing shortness of breath, fatigue, angina, or palpitations - the ailments may also occur at rest.

The introduction of numbers to indicate the consecutive stages of the advancement of coronary heart disease (CCS) or heart failure (NYHA) often leads to errors and mistakes. Researchers use statistical analysis methods which are not correct in such cases, and on this basis they often draw erroneous conclusions concerning the conducted research.

The methods used in publications [5] or [6] constitute examples of such incorrectly conducted analyses (Tables I, II).

In both these works, average values were established within each scale after implementing the CCS and NYHA scales and providing the number of patients assigned to each of the 4 classes in the CCS and NYHA scales based on the patients' clinical symptoms. It is not clear how these average values should be interpreted. What is the meaning of an average value of 2.93 ± 0.67 – Table I work [5] – given that the NYHA scale is composed of 4 separate classes to which patients are classified on the basis of the established clinical symptoms? Which of these symptoms should be assigned to the patients with disease advancement of 2.93 in comparison to classes II or III? Additionally, the authors decided to provide the standard deviations for the average values that they established. Average values and standard deviations are mathematically valid for continuous variables with normal distribution. In the case of the NYHA scale, the variables are qualitative and one is not allowed to calculate average values for them, not to mention their standard deviation. Thus, a question arises concerning the way in which the authors would conduct their statistical analysis if the NYHA or CCS scales had classes A, B, C, and D instead of I, II, III, and IV respectively. Would it still be possible to calculate an average NYHA class in the studied group? When conducting analogous calculations (Equation 1):

$$\bar{X}_{NYHA^*} = \frac{176 \times 2 + 370 \times 3 + 129 \times 4}{675} = 2.93 \quad \text{Equation 1}$$

we end up with an expression whose value is not possible to establish (Equation 2).

$$\bar{X}_{NYHA} = \frac{176 \times B + 370 \times C + 129 \times D}{675} \quad \text{Equation 2}$$

A correctly performed analysis of the data from, for example, Table IV in work [5] should contain the number of patients for the individual subgroups of the NYHA classification, as done in Table I of work [5]. This would allow for adequate comparison between groups of patients, depending on classification (Table IV qualified/non-qualified), using the χ^2 test or the χ^2 test with appropriate amendments depending on the numbers of patients in groups II, III, and IV [9]. The authors, instead of providing the numbers of patients for the individual subclasses, incorrectly averaged the results and provided the values of standard deviations for the qualified and non-qualified patients. It is a well-known fact that it is not permissible to average

Tab. I. Fragment of Table I [5]. Description of patients during OHT qualification

Variable	Average ± SD	N
NYHA*(class)	2.93 ±8.1	675
II	26%	176
III	55%	370
IV	19%	129

Table II. Fragment of Table I. [5] Classes marked B, C, D, respectively. Description of patients during OHT qualification

Variable	Average ± SD	N
NYHA*(class)	2.93 ±8.1	675
B	26%	176
C	55%	370
D	19%	129

values within a classification, not to mention comparing these averages using methods of statistical analysis and formulating conclusions on this basis.

As already stated, the CCS and NYHA scales are ordinal scales with four classes each. Therefore, within each of these classes it is possible to test hypotheses concerning trend occurrence. On this basis, it is possible to assess whether there exists a linear relationship between the proportions in individual subclasses. In order to conduct such analysis, it is possible to use, for example, the Cochran-Armitage test [10], which may be found in the Statistica, SAS, or MedCalc packages. Moreover, splitting patients into NYHA or CCS subgroups allows for the study of the occurrence of statistically significant differences between continuous variables within different subclasses of the analyzed class. For example, it is possible to check whether there are statistically significant differences in the values of systolic and diastolic pressure or in the left ventricular end diastolic diameter (LVEDD) between patients of different NYHA subclasses. In this case, if we wanted to compare the averages of two subclasses, we would have to use the Student's parametric *t*-test for the normal distribution of continuous variables and similar numbers of patients in subgroups. If the distribution is not normal, we would use the non-parametric Mann-Whitney U test.

To sum up, it needs to be recognized that numbers do not always have direct value references, which the authors of the NYHA scale emphasized by the use of Roman numerals. Averaging data of this type is mathematically erroneous and has no medical meaning. Presenting the percentage of the analyzed patients in the individual classes and conduct-

ing analysis using the χ^2 test or Fisher's exact test may serve as a solution in such cases. Therefore, when the analyzed numbers can be substituted with letters without any loss of information, one should remain cautious, because certain mathematical operations may not be applicable.

References

1. Fisher JD. New York Heart Association Classification. *Arch Intern Med* 1972; 129: 836.
2. Campeau L. Letter: Grading of angina pectoris. *Circulation* 1976; 54: 522-523.
3. Killip T 3rd, Kimball JT. Treatment of myocardial infarction in a coronary care unit. A two year experience with 250 patients. *Am J Cardiol* 1967; 20: 457-464.
4. Forrester JS, Diamond GA, Swan HJ. Correlative classification of clinical and hemodynamic function after acute myocardial infarction. *Am J Cardiol* 1977; 39: 137-145.
5. Korewecki J, Browarek A, Zieliński T, Sobiećzańska-Matek M, Piotrowska M, Zembala M, Zakliczyński M, Rozentryt P, Barańska-Kosakowska A, Sadowski J, Przybyłowski P, Maliniak I, Garlicki M. Rokowania chorych z ciężką niewydolnością serca, wstępnie kwalifikowanych do przeszczepu serca – na podstawie ogólnopolskiego rejestru POLKARD 2003–2005. *Kardiologia i Torakochirurgia Polska* 2006; 3: 308-322.
6. Śpiewak M, Głowczyńska R, Matek A, Grabowski M, Filipiak KJ. Rozpoznanie i klasyfikacja stabilnej choroby wieńcowej. *Przew Lek* 2006; (6): 34-41.
7. Matek Ł, Chojnowska L, Kłopotowski M, Dąbrowski M, Mączczyńska R, Demkow M, Witkowski A, Kuśmierczyk B, Piotrowicz E, Konka M, Rużyłło M. Long-term follow-up of patients with hypertrophic obstructive cardiomyopathy treated with percutaneous alcohol septal ablation. *Postępy Kardiologii* 2009; 5: 167-171.
8. Deja M, Janusiewicz P, Biernat J, Malinowski M, Gołba K, Domaradzki W, Bachowski R, Jasiński M, Ceglarek W, Woś S. Wyniki chirurgicznej rekonstrukcji lewej komory serca. *Kardiologia i Torakochirurgia Polska* 2008; 5: 112-115.
9. Kwasiborski PJ, Maria Sobol M. Test niezależności chi-kwadrat i jego zastosowanie w interpretacji wyników badań klinicznych. *Kardiologia i Torakochirurgia Polska* 2011; 4: 550-554.
10. Armitage P. Tests for linear trends in proportions and frequencies. *Biometrics* 1955; 11: 375-386.