Original article

Descriptive presentation of wound care data

Keywords descriptive statistics, categorical, numerical, repeated measurements

For referencing Stephenson J. Descriptive presentation of wound care data. WCET® Journal 2022;42(3):30-33

DOI https://doi.org/10.33235/wcet.42.3.30-33

Submitted 1 August 2022, Accepted 1 September 2022

INTRODUCTION

Most published quantitative research in the field of wound care will include elements of both descriptive and inferential statistics. Descriptive statistics, which normally precede the presentation of inferential tests, describe a study sample, using summary statistics, tables and graphs. No inference is involved. Inferential statistics, which includes significance testing and confidence intervals, is concerned with the inferences made from sample data to a wider parent population, and is not the subject of this editorial.

The aim of descriptive statistical analysis is to condense data in a meaningful way and extract useful information from it. Data may take various different forms, of which the distinction between two forms, *categorical* and *numerical*, is important for decision-making concerning the most appropriate method needed to provide an effective descriptive summary of the data. Categorical variables are sometimes further sub-divided into *nominal* variables (i.e. those where there is no underlying ordering to the categories) and *ordinal* variables (with some underlying order). The categories themselves are often termed *levels*.

In most wound care studies, the most common sources of data are probably the wound itself and the patient with the wound. An example of patient-level categorical data is patient sex (levels: male and female); an example of wound-level categorical data is tissue type (levels: slough, necrotic etc.) An example of patient-level numerical data is patient age in years; an example of wound-level numerical data is wound length. We may also collect and report data at the aggregate level; for example, the proportion of patients with a healed wound by 30 days, or the mean number of patients treated per month by a clinical team.

Sometimes the distinction between categorical and numerical data is not clear. Responses from questionnaire items, such as the commonly-encountered 5-point Likert questionnaire item, are, strictly speaking, ordinal, but are often treated as numerical, particularly when dealing with a score which is a sum of multiple items. Other types of data can be formulated as either categorical (e.g. the proportion of wounds healed within 30 days) or numerical (e.g. the number of days to healing), depending on the context and the aims of the study.

John Stephenson
PHD FRSS(GradStat) CMath(MIMA)
Senior Lecturer in Biomedical Statistics
University of Huddersfield, United Kingdom
Email J.Stephenson@hud.ac.uk

PRESENTING DESCRIPTIVE DATA IN TEXT AND TABLES

Many wound care studies generate far too much data to present it all in text. Often only key results are presented in text, with the bulk of the data appearing in tabulated form, possibly in an appendix. Whether in text or in tabulated form, standard presentation for a numerical variable is a measure of average, followed by a measure of dispersion (i.e. spread) in brackets. The measure of average quoted is almost always the mean (i.e. arithmetic mean) or the median. Medians, which are not distorted by outlying values, are usually preferred when data is likely to be skewed – such as time to wound healing or some other event, or when we are dealing with ordinal quantities (such as the sum of Likert-style questionnaire items) which are assumed to be equivalent to numerical data – otherwise, the mean, which uses all the data values, is generally preferred.

The measure of dispersion quoted is usually either the standard deviation (commonly abbreviated to SD) or the range and/or inter-quartile range (commonly abbreviated to IQR). The range of a data set is easy to calculate (simply the difference between the two extreme values) but is based only on those two measures, disregarding all others. It is distorted by outliers. The IQR, which is calculated as the range from the 25th to the 75th percentile of the data, is more robust to distortion, but still does not take into account much of the data set.

By contrast, the SD uses every observation, but can be sensitive to outliers and is generally inappropriate for skewed data. It also has the advantage that it is always in the same units of measurement as the raw data, which can help with interpretation; in normally distributed data, approximately two thirds of all observations will lie within one standard deviation of the mean. So, for example, if we are told that the mean wound diameter in a large study of venous leg ulcers is 20mm, with an SD of 4mm, then if the data is normally distributed, we can infer that about two-thirds of wounds have a diameter between 16mm (1 SD below the mean) and 24mm (1 SD above the mean). The remaining one-third of wounds would be expected to be relative outliers, either below 16mm or above 24mm in diameter.

Common pairings for presenting descriptive data are mean and SD, median and range, and median and IQR. Other measures of average and spread, such as the geometric mean, mode and mid-range, are much less commonly encountered.

Standard presentation for a categorical variable is frequency, plus percentage and/or proportion. Generally valid percentages are quoted, disregarding invalid or missing data. For example, an audit of pressure injuries in a particular hospital ward ICU might record a number of Stage 1, 2 and 3 pressure injuries in ICU patients, but some patients on the ward are missed out from the audit. It would probably be more appropriate to quote the numbers of patients with a Stage 1 pressure injury as a proportion (and/or percentage) of the patients who were actually audited, not as a proportion of all patients.

Table 1, adapted from Ousey et al.[1], shows an example of tabulated data in a quite typical format. It includes both a numerical variable (age), summarised using mean and SD in each study group, and several categorical variables, summarised using frequency and valid percentage. Here the proportion is also given. The levels of each categorical variable considered are indented below the name of the variable itself. This amount of data would be difficult to absorb in text, and the side-by-side format of the table facilitates an easy comparison of group characteristics that would not be so apparent in data presented in text.

Table 1. Example of tabulated data [adapted[1]]

Characteristic	**Mattress type**	
	Pressure re-distributing	**Standard**
Patient age in years (mean (SD))	73.0 (18.5)	76.6 (10.1)
Patient gender		
Male	17/23 (73.9%)	16/28 (57.1%)
Female	6/23 (26.1%)	12/28 (42.9%)
Risk of pressure injury (Waterlow score)		
Low risk (<10)	0/16 (0.0%)	1/16 (6.3%)
At risk (10–14)	13/16 (81.3%)	15/16 (93.8%)
High risk (15–19)	1/16 (6.3%)	0/16 (0.0%)
Very high risk (20+)	2/16 (12.5%)	0/16 (0.0%)
Pressure injury present		
Yes	4/24 (16.7%)	7/27 (25.9%)
No	20/24 (83.3%)	20/27 (74.1%)
Skin temperature control		
Very good or excellent	8/18 (44.4%)	5/13 (38.5%)
Good, adequate or poor	10/18 (55.6%)	8/13 (61.5%)

Note that the denominator is different for the different patient characteristics featured in the table; not all characteristics will have been reported on all patients. The levels of the *Skin temperature control* variable have been 'condensed' from five individual categories into two contrasting levels; this is a common device when data is spread too thinly across multiple levels for meaningful analysis, or when highlighting a contrast between two meaningful clinical states. The Waterlow variable has been transformed from its original numerical scale into an ordinal categorical variable; at the cost of a certain loss of information, this also allows comparison across levels of risk in common clinical use.

PRESENTING DESCRIPTIVE DATA IN GRAPHICAL FORM

Many different types of graphs are available, and most can be produced easily using modern software. Not all graphs are suitable for all types of data, however. Pie charts and bar charts are both designed to visually illustrate the relative frequencies of multiple levels of categorical variables. Despite its ubiquity, the pie chart does not seem to offer anything that a bar does not; most people find it harder to assess the relative size of sectors of a circle than they do of the heights of columns. Neither representation works well to display very large numbers of categories (which are hard to compare visually).

The bar chart can also be used to represent a quantity expressed as a proportion – Ousey et al.[2] presented the proportion of patients with pressure ulceration pre- and post-implementation of a pressure reduction implementation programme in terms of a simple bar chart (Figure 1). The 'whiskers' around the bars represent confidence intervals, a measure of uncertainty in the quantity being measured.

A useful extension to the bar chart is the clustered bar chart which allows display of two factors concurrently. Figure 2 is a neat representation of the interplay between two categorical factors – pressure grading system status (with levels represented by the left- and right-hand clusters) and policy of referral (bars within a cluster).

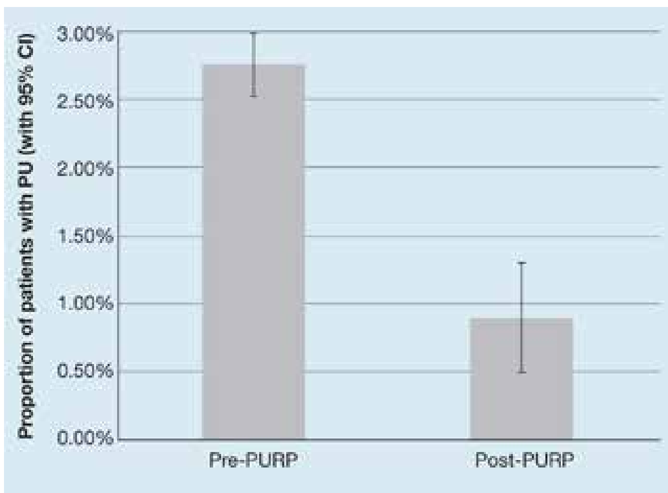

Figure 1. Example of a simple bar chart [adapted[2]]

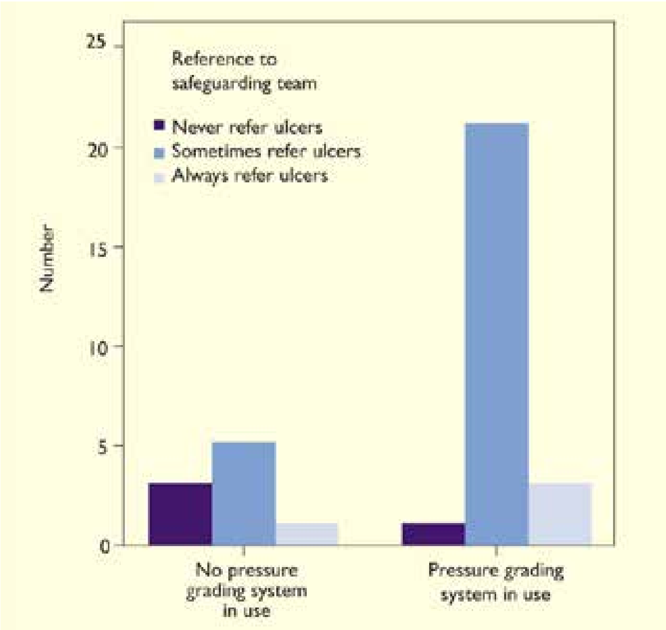

Figure 2. Example of a clustered bar chart [adapted[3]]

Different representations are required for numerical data. For example, a histogram, which is often confused with a bar chart, was used by Barakat-Johnson et al.[4] to represent the time of response to communication with a wound specialist reported by patients using a digital app (Figure 3). It can be distinguished from a bar chart by the lack of gaps between the bars, reflecting the representation of a continuous measure rather than distinct categories. This sort of data can also be represented using a box plot, although box plots do not provide information about the full distribution of a data set.

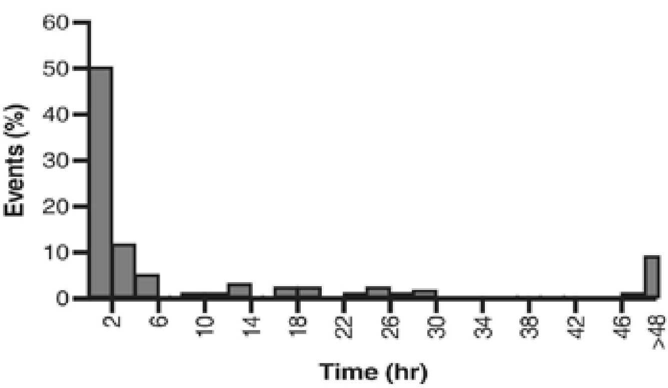

Figure 3. Example of a histogram [adapted[4]]

Mixed representation

The relationship between a numerical variable (such as peak pressure index) and categorical variables (such as BMI category and body position) can be neatly combined in a single representation using a line graph, as reported by Coyer et al.[5]. Figure 4 shows body position distinguished by the colour and shading of the line, and BMI category by the position on the x-axis.

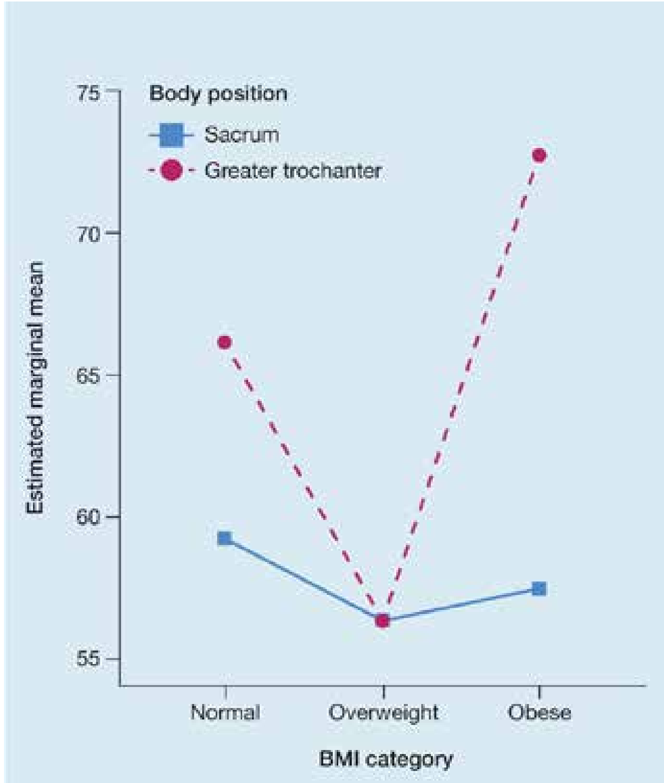

Figure 4. Example of a line graph [adapted[5]]

The key take-home message from this graph is that the factors interact – the effect on peak pressure of BMI category depends on where it is measured. This effect would not be immediately apparent had the same data been presented in tabulated form.

Repeated measures

Many wound care studies lend themselves to repeated measurements, for example, to examine the healing trajectory of a wound by monitoring its length at weekly intervals until healing, or investigating trends by audits of institutional aggregated data. Stephenson et al.[6] presented longitudinal data (here, the number of observations of category 2 pressure injuries reported in a health organisation at monthly intervals over a period of several years) as a line graph (Figure 5), with the dotted line illustrating the underlying smoothed time-dependent trend. Here the graph is illustrating seasonal trends, an overall year-on-year downward trend, and the relationship between one data point and the preceding point (auto correlation) – effects that would be almost impossible to discern from tabulated data alone.

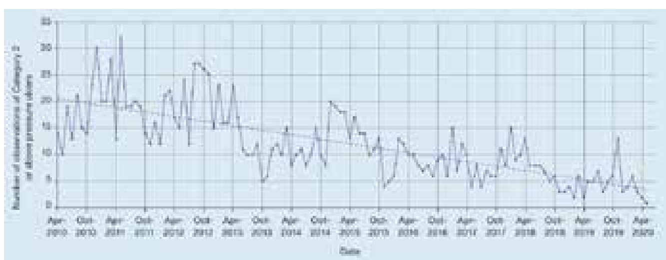

Figure 5. Example of a line graph showing longitudinal data [adapted[6]]

TABLES OR GRAPHS?

It is not always easy to decide whether a table, graph or both is needed to summarise data. Graphs show trends and patterns in data, and the relationship between one variable and another, that would not necessarily be apparent in the same data presented in tabulated form. Tables give values to a level of precision that is generally unavailable in most graphical presentations.

CONCLUSIONS

Effective presentation of descriptive wound care data can allow a reader to quickly absorb trends and patterns in data, to compare group characteristics, and to assess the magnitude of effects. The questions that arise from wound care studies – in which we may be looking to compare the benefit of one treatment against another, maybe examine the change in wound parameters over time, or simply summarise the extent of wounds in an audit study – can often be answered simply and effectively using descriptive analysis, although, usually, such an analysis would be followed up with an inferential assessment. However, the ease with which modern software can generate graphs of any kind can sometimes be a barrier to effective communication. Many published examples exist of graphs that add little or nothing to understanding and need to be used with care.

While descriptive statistics do not facilitate drawing conclusions beyond available data or reject any study hypotheses, they can be a valuable way of adding insight to a study and require little or no specialist statistical knowledge to understand.

REFERENCES

1. Ousey K, Stephenson J, Fleming L. Evaluating the Trezzo range of static foam surfaces: results of a comparative study. Wounds UK 2016;12(4):66–73. ISSN 1746-6814.
2. Ousey K, Stephenson J, Blackburn J. Sub-epidermal moisture assessment as an adjunct to visual assessment in the reduction of pressure ulcer incidence. J Wound Care 2022;31(3).
3. Ousey K, Kaye V, McCormick K, Stephenson J. Investigating staff knowledge of safeguarding and pressure ulcers in care homes. J Wound Care 2016;25(1).
4. Barakat-Johnson et al. The viability and acceptability of a Virtual Wound Care Command Centre in Australia. Int Wound J 2022;1–17.
5. Coyer F, Clark M, Slattery P, Thomas P, McNamara G, Edwards C, Ingleman J, Stephenson J, Ousey K. Exploring pressures, tissue reperfusion and body positioning: a pilot evaluation. J Wound Care 2017;26(10).
6. Stephenson J, Ousey K, Blackburn J, Javid F. Using past performance to improve future clinical outcomes in pressure ulcer prevention. J Wound Care 2021;30(6).