Radiography (2010) 16, e1—e7

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/radi

ELSEVIER

THE SOCIETY AND COLLEGE OF RADIOGRAPHERS

REVIEW ARTICLE

An introduction to descriptive statistics: A review and practical guide

Gill Marshall*, Leon Jonker

Faculty of Health, Medical Sciences and Social Care, University of Cumbria, Bowerham Road, Lancaster LA1 3JD, UK

Received 28 May 2009; revised 27 November 2009; accepted 3 January 2010
Available online 6 February 2010

KEYWORDS
Descriptive statistics;
Levels of measurement;
Measures of central tendency;
Dispersion;
Normal distribution;
Standard deviation

Abstract This paper, the first of two, demonstrates why it is necessary for radiographers to understand basic statistical concepts both to assimilate the work of others and also in their own research work. As the emphasis on evidence-based practice increases, it will become more pressing for radiographers to be able to dissect other people's research and to contribute to research themselves. The different types of data that one can come across are covered here, as well as different ways to describe data. Furthermore, the statistical terminology and methods used that comprise descriptive statistics are explained, including levels of measurement, measures of central tendency (average), and dispersion (spread) and the concept of normal distribution.

This paper reviews relevant literature, provides a checklist of points to consider before progressing with the application of appropriate statistical methods to a data set, and provides a glossary of relevant terms for reference.

Introduction

Difficulty in understanding statistics is one of the most frequently reported barriers stopping nurses applying research results to their practice, thereby limiting the opportunity to undertake practice which is based on research evidence.[1–3] This is also likely to apply in other health care professions, such as radiography. In fact, the authors of this review started out not having a formal specialist degree in statistics. The intention here is to demonstrate that an appreciation of statistics can be gained through self-teaching and by accessing literature that is aimed at novice researchers or clinical staff and allied health professionals who perhaps never have been exposed to in-depth statistical tuition during their formal education, to thereby better understand statistics. To further help the reader on his/her way, the authors will introduce relevant statistical terminology and the basic elements of the application of statistics. The authors wish to stress that with more complex statistical tasks one should always consider contacting a qualified statistician for support and feedback, preferably before

* Corresponding author. Tel.: +44 1524 384384x2246; fax: +44 1524 384385.
E-mail address: gill.marshall@cumbria.ac.uk (G. Marshall).

1078-8174/$ - see front matter
doi:10.1016/j.radi.2010.01.001

starting data collection to ensure that the data to be gathered is appropriate.

Statistics in radiography

The measurement of parameters in medical care is one of the first points for facilitation of quality improvement.[4] Whilst it is certainly not necessary for radiographers to be full qualified statisticians, a working knowledge of statistics is desirable in order to be able to assess and understand the implications of research for radiographers to evolve their practice. Furthermore, if a radiographer is undertaking a research project, he/she must be statistically literate to interact with other researchers and statisticians and in order to ensure a project has the best methodological design which may include selecting appropriate statistical methodology which will test a predetermined hypothesis; moreover, this test should then be undertaken appropriately.[1]

The quality of statistics education provided to students whilst they are on a course or taking a degree is reported in the literature to be inadequate. Furthermore, students soon forget what has been taught to them, particularly if it is not applied.[5,6] These two scenarios form a barrier to the understanding of research articles. Informing someone on how to read and dissect statistics is found to be most effective when the statistical concepts are taught first, before commencing the teaching of ''how to undertake'' the statistics.[7] As radiography strives to become an autonomous profession it must ''carve out a knowledge base that is dynamic and forward thinking''[8] and this clearly needs cognisance of statistics read by radiographers in articles or applied to their own studies. It is claimed that ''today's medicine is only as good as the research fifteen years ago'', showing the considerable time taken between the findings of research being reported and them filtering into clinical practice.[6] Although the aim of evidence-based practice is to narrow the theory—practice gap through evidence from research informing practice, it appears that the research is perhaps sometimes not presented in a way which is user-friendly and comprehensible to practitioners, who consequently will not or cannot apply it to their daily practice.[9] It is thus incumbent on researchers to think about their audience, when publishing, so that their results are clinically relevant and meaningful.

In a study by Welch and Gabbe in 1996, 32% of the articles were either devoid of statistics or claimed significance in the results without the use of statistics.[9] A reader with only a knowledge of descriptive statistics will only typically understand 44.5% of the data in articles, whilst those with a knowledge of common statistical tests will increase the access rate of understanding data produced in articles to 80.5%.[9]

Types of data

Statistics are used to demonstrate the meaning of the data, and are based on numbers, e.g. the patient's heart rate, or assigned numbers to qualitative attributes such as eye colour. Statistics can be used descriptively to illustrate the characteristics of a group of observations i.e. the raw data; this is called descriptive statistics.[10] In statistical terms, there are different levels of measurement for scoring variables. There are two main categories:

- Categorical data
- Continuous data

Categorical data can be divided further into four subcategories:

- Binary data

In such instances, only two outcomes or measurements are possible. For instance, if the patients' survival is recorded there can only be two outcomes — survival or death. Another example is whether or not a radiological report has been filled in correctly. Such data is usually summarised using proportions or odds.

- Nominal data

Whilst this is the least robust type of data for categorising data into mutually exclusive categories, without hierarchical ranking, this is nonetheless a much used and valuable type of data. An example could be blood groups, i.e. O, A, B, AB, or distinguishing groups of professions e.g. radiographers, radiologists and nurses. Both binary data and nominal data can be represented or stored by allocating numbers to categories (radiographers = 1, radiologists = 2 and nurses = 3, where the numbers have no numerical significance).[11] With this data the measure of central tendency is the category with the most cases, known as the mode.

- Ordinal data

Here the data has a clear order or hierarchy but not on a calibrated scale e.g. strongly agree, agree etc., from a so-called Likert scale with a statement provided.[12] An example is the level of pain caused by mammography e.g. no pain, mild discomfort, moderate pain, severe pain. The categories are mutually exclusive and, as above, the numeric values attributed to the categories are not absolute measurements but they do order the data. In ordinal data, the measure of central tendency is called the median and the category which is the middle of the rank-ordered description is thus referred to as the median.

Binary, nominal and ordinal scales are considered *discrete variables* because the data is classified into discrete non-overlapping variables.[13]

- Interval/ratio data

This is the strongest type of data, with ratio data being stronger data than interval data, as ratio data has a true zero value, whilst interval data does not.[14] Such data is achieved by the use of a calibrated scale to provide quantitative measurements e.g. density readings from a densitometer, weight in kilograms, or blood pressure. Such data can be plotted into a histogram. Ratio or interval data may be summarised by the mean or the median (as a measure of central tendency) depending on the distribution. To illustrate the point, an example of interval data would be an IQ test score. Ratios of these measurements cannot be applied — an IQ of 140 in subject 1 versus 70 in subject 2 does not mean subject 1 is twice as clever as subject 2. On the other hand, examples of ratio data show that they all have a constant scale which includes a zero e.g. two hundred metres is twice as long as one hundred metres. Such data can be plotted into a histogram and the midpoint of the curve will again be the measure of central tendency. If a variable is normally distributed (i.e. it produces a symmetrical bell curve) the mean, median and the mode value will be approximately equal. If the researcher is going on to undertake inferential statistics i.e. tests that allow the researcher to draw inferences from the data (covered in a subsequent article) it must be decided whether the data is normally distributed. If it is then the mean and standard deviation will suffice as summary statistics. If the data is not normally distributed then the five order statistics are used i.e. minimum, first quartile, median, third quartile and maximum. The range and the inter-quartile range can easily be derived from these but are not usually used explicitly as summary statistics. The outcome of testing this will determine whether the researcher can use parametric statistics or non-parametric statistics. For interval/ratio data from a continuous scale the range, inter-quartile range and standard deviation are used to report the spread or width of the data. Data from interval or ratio scales are described as *continuous data* and thus provides *continuous variables* because the data represents an underlying continuum where there are potentially an infinite number of values.

Distribution of data

This article focuses on the statistical methodology that can be applied to descriptively define the data. These methods are numerical procedures or graphical techniques e.g. bar charts, histograms, frequency polygons and pie charts, used to organise, present and describe the characteristics of a sample e.g. they provide a summary measures of the characteristics. Descriptive statistics seek to describe the midpoint of a spread of scores, called the measure of central tendency, and the spread of scores which is called the dispersion, of which variance is an example.[14] In order to understand this, it is necessary to consider the levels of measurement, because certain (subsequent) tests only work with appropriate levels of measurement. If measurements are taken from a large random sample e.g. of the weight of adult patients having contrast enhanced CT and a frequency polygon is plotted of the results, it is likely a bell shaped curve is produced which shows that the variables of a sample are normally distributed. This bell shape is called a normal or Gaussian distribution (See Figure 1). The word 'normal' here means that the data complies with a distribution pattern that mathematically allows parametric statistical tests to be applied.

A hypothetical measurement of an artery of 100 patients, presented through a histogram, with Gaussian bell curve highlighting the normal distribution pattern.

In radiography normal distribution of measurements may be seen when plotting the sizes and volumes of certain

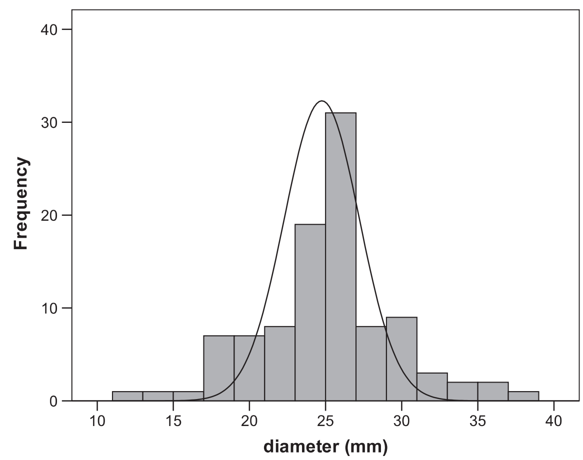

Figure 1 Example of data with a normal distribution.

anatomical structures, such as the eyeball, optic nerve or as shown in Figure 1, vascular structures. Generally speaking, the measurements for such structures do not differ much between people. However, often a normal distribution is not achieved; outliers can be common which will give the bell curve a 'tail' at either the left hand or right hand side. Examples here would be a histogram of patients' weight, where the distribution curve would be skewed towards a higher weight (the curve would be described as positively skewed) — see also Figure 2.

The same data as represented in the histograms of Figure 1 is summarised.

If data does not exhibit a normal distribution, for example when there are some outliers, the presentation of data can be altered to reach a normal distribution. A common mathematical procedure to carry out is transforming data logarithmically; it does not change the actual data but it does the way it is distributed on a graph and thereby facilitates analysis of the data as if it were normally distributed. If normalisation is not possible, appropriate non-parametric tests should be applied to the data. Non-parametric tests take into account the fact that the data is skewed or has various outliers and are therefore more conservative. The statisticians Altman and Bland illustrate these points very clearly in a series of short articles that they have published over the years in the British Medical Journal.[15–18]

Standard deviation is a measure of how spread out the data is, the variance of it. The narrower the standard deviation the closer to the midpoint of the data all results will be.[17] Without quoting the mathematical formula behind it, the standard deviation allows for expressing variance using the same units as those used for the observations or measurements. Generally speaking approximately 2/3 of all observations or measurements lie within one standard deviation of the mean (the top of the distribution graph in Figure 1), and 95% lie within two standard deviations of the mean. The inter-quartile range is becoming more common in reporting of descriptive statistics from continuous data. This statistic represents the middle 50% of the sample showing its dispersion, and is not influenced by outliers.[18] Table 1 contains examples of the standard deviation for the two example data sets as well as

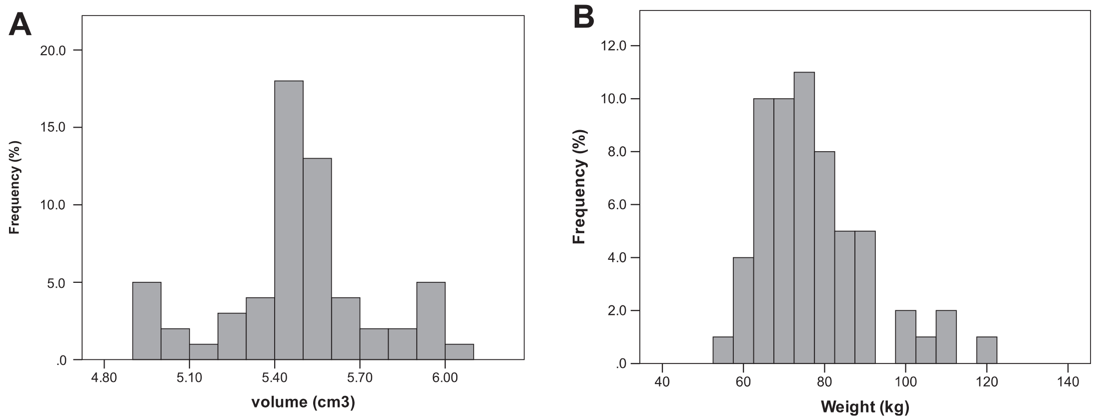

Figure 2 Examples of data with normal and skewed distribution. These are distributions of hypothetical data for 60 patients concerning the volume of an eyeball with normal distribution (A) and bodyweight with negatively skewed distribution (B), respectively.

the inter-quartile range. From the standard deviation particularly it can be concluded that for the bodyweight data more measurement points lie further away from the mean.

Collection and presentation of data

It is essential to clarify the appropriate type of data needed to answer the research question at the design stage of a research project, so it can be gathered. The level of measurement needs to be identified, allowing identification of the statistical procedure to be used and decision making on the sample size. Descriptive statistics are the easiest to undertake and interpret and they are a useful way to summarise data and provide a description of the sample. Figure 3 provides another example of how to summarise data; in this instance the hypothetical ratio between two measurements of the diameter of the bronchial and pulmonary arteries for 100 patients. Box plots allow a clear overview of the characteristics of a data set, as shown in Figure 3. The box represents the 25th and 75th percentile values, whereas the bar inside the box represents the 50th percentile (median). The 'whiskers' attached to the box represent the lowest and highest values (the range) of the data, bar extreme outliers or extreme values.

Ultimately, descriptive statistics cannot be used to demonstrate causal analysis which requires the use of inferential statistics, which allow us to generalise from a sample to a larger population. Data analysis can be taken one step further by inferring from the sample group

Table 1 Descriptive statistics of two example sets of data.

	Volume (cm^3)	Weight (kg)
Mean	5.50	76.75
Median	5.50	73.00
Standard deviation (% of mean)	0.27 (5%)	13.34 (17%)
Inter-quartile range (25%–75%)	5.38–5.60	67.25–83.00

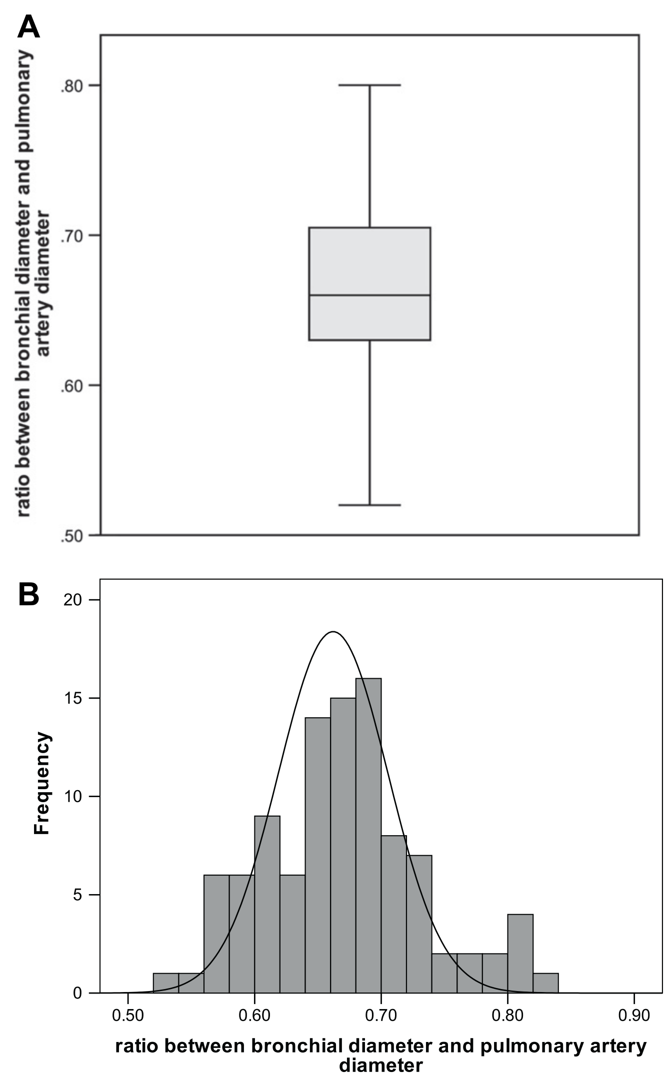

Figure 3 Presentation of the ratio between bronchial and pulmonary artery diameters in histogram (A) and box plot (B) format.

generalisations that can be applied to a wider population which is called inferential statistics — these are considered in our second article.

The understanding of statistics is not just important to the researcher but readers of research output (manuscripts, reports, theses) need to be cognisant of the level at which variables are measured, the statistical tests used and they must be aware that statistical procedural error by researchers is not uncommon.[9] Typically this error is not in the calculation of the test result, which nowadays is rarely undertaken by hand but using statistical packages such as Excel, Minitab, Supastat, and the comprehensive and commonly used SPSS (Statistical Package for Social Science), but in the test selection. Some of the above mentioned statistical packages can be freely downloaded, albeit in some cases just for a limited period.

Checklist for understanding and applying descriptive statistics:

- Regarding your data do you just want to describe and summarise it?
  - o If so only descriptive statistics will suffice.
- What measure of central tendency is appropriate to the data?
- Are your variables discrete or continuous?
- Is your data nominal, ordinal or ratio/interval data?
  - o Once you have completed descriptive statistics, can more information be drawn out of the data?
  - o If so consider the use of inferential statistics which allows drawing conclusions from the data.

Conclusion

Statistics provide a way of describing collected numerical data that can be assimilated by readers, allowing the research outcome to be used for evidence-based practice and thereby narrowing the theory—practice gap. Descriptive statistics (which in certain cases can be labelled 'explorative statistics') is perfectly suited for collating and summarising quantitative data. It allows a radiographer to gain an insight in, for example, potential trends in the demographics of certain patient groups or to detect certain trends in treatment outcomes through observational studies. Such observations can then be used as a starting point to further investigate the reasons behind these characteristics and trends. The next step will then be to apply inferential statistics to allow drawing conclusions applicable to a larger group than the sample used to generate the data. Statistics can be difficult to apply and understand, especially to those without experience. It is essential that the type of data collected and its analysis or representation is appropriate, so that the research question can be answered meaningfully and so that information can be obtained from the collected data. It is important to consider what type of data will be collected and presented as soon as the research question is identified.

Acknowledgement

I would like to thank Sue Griffiths for her proof reading of this article.

References

1. Rickard C. Statistics for clinical nursing practice: an introduction. *Australian Critical Care: Official Journal of the Confederation of Australian Critical Care Nurses* 2008;**21**:216—9.
2. Justham D, Timmons S. An evaluation of using a web-based statistics test to teach statistics to post-registration nursing students. *Nurse Education Today* 2005;**25**:156—63.
3. Davidoff F, Haynes B, Sackett D, Smith R. Evidence-based medicine: what it is and what it isn't. *British Medical Journal* 1996;**312**:71—2.
4. Allison JJ, Calhoun JW, Wall TC, Spettell CM, Fargason Jr CA, Weissman NW, et al. Optimal reporting of health care process measures: inferential statistics as help or hindrance. *Managed Care Quarterly* 2000;**8**:1—10.
5. Reznick RK, Dawson-Saunders E, Folse JR. A rationale for teaching of statistics to surgical residents. *Surgery* 1987;**101**:611—7.
6. Bolton J. Clinicians and the ''S — word''. *Clinical Chiropractic* 2006;**9**:88—91.
7. Hong E, O'Neil H. Instructional strategies to help learners builds relevant mental models of inferential statistics. *Journal of Educational Psychology* 1992;**84**:150—9.
8. Manning D, Hogg P. Writing for publication. *Radiography* 2006;**12**:77—8.
9. Welch G, Gabe S. Review of statistics usage in the American journal of gynaecology. *American Journal of Obstetrics and Gynecology* 1997;**176**(5):1138—41.
10. Parahoo K. *Nursing research principles, processes and issues*. 2nd ed. Basingstoke: Palgrave Macmillan; 2006. p. 468.
11. Botti M, Endacott R. Clinical research 5: quantitative data collection and analysis. *Intensive and Critical Care Nursing* 2005;**21**(3):187—93.
12. Likert R. A technique for the measurement of attitudes. *Archives of Psychology* 1932;**140**:1—55.
13. Polgar S, Thomas S. *Introduction to research in health sciences*. 3rd ed. Melbourne: Churchill Livingstone; 1995. pp. 243—306.
14. Fisher M, Marshall A. Understanding descriptive statistics. *Australian Critical Care* 2009;**22**:93—7.
15. Altman D, Bland J. The normal distribution. *British Medical Journal* 1995;**310**:298.
16. Bland J, Altman D. Transformation, means, and confidence intervals. *British Medical Journal* 1996;**312**:1079.
17. Altman D, Bland J. Standard deviations and standard errors. *British Medical Journal* 2005;**331**:903.
18. Altman D, Bland J. Quartiles, quintiles, centiles and other qualities. *British Medical Journal* 1994;**309**:996.

Glossary

Bar charts: these are used where the variable plotted along the horizontal axis is nominal or ordinal i.e. the categories shown along this axis have no quantitative or numerical value (nominal), or clear order or hierarchy but not on a calibrated scale (ordinal). The vertical bars are drawn separately e.g. the frequency distribution of the number of qualified radiographers in the UK categorised by age group of the radiographers.

Cohort: a defined group of people e.g. all women in the Morecambe bay area who have been screened for breast cancer in the last year.

Continuous variables: the data represents an underlying continuum where there are potentially an infinite number of values.

Table 2 Common statistical terms and their definitions.

Measures of central tendency	
Term:	Definition:
Mode	The numerical value with the greatest frequency
Median	The middle score of a rank-ordered distribution
Mean	The average score
Dispersion	
Frequency distribution	The number of cases per category
Range	The distance between the highest and lowest score
Inter-quartile range	The range within which the middle 50% of the scores fall
Standard deviation	The root-mean-square deviation from the mean
Variance	The square of the standard deviation

Descriptive statistics for each level of measurement.[14]

Data: numbers or values collected as a result of measurements. They could be counts or frequencies or actual numerical values or scores.

Descriptive statistics: Statistics that can be used descriptively to illustrate the characteristics of a group of observations i.e. the raw data.

Discrete variables: when the data is classified into discrete not overlapping variables.

Frequency polygon: this is an alternative to a histogram and is preferred when the variable is continuous. It is produced by placing dots at the tops of the centres of the bars of the equivalent histogram, and then joining the dots with straight lines. It is often used to compare two frequency distributions.; When measurements are taken from a large random sample and a frequency polygon is produced the shape of the distribution is always the same. This is a bell shaped curve called a normal or Gaussian distribution.

Histograms: here the categories are measured on a numerical scale. The vertical bars are drawn touching e.g. the frequency distribution of the number of children per family living in Lancaster. Histograms can be used for discrete or continuous data. If used for continuous data, it is partitioned into ranges, the size of which depends on the sample size to enable an optimum number of bars to be displayed.

Inferential statistics: statistics that can be used to infer from the sample group generalisations that can be applied to a wider population.

Interval data: is stronger data than nominal or ordinal data. It can achieved by the use of a calibrated scale to provide quantitative measurements. The difference between interval and ratio data is that ratio data has a true zero, thus interval data is weaker data than ratio data.

Likert scale: is a scale commonly used in questionnaires, and is the most widely used scale in survey research. It is used for ordered category data. When responding to a Likert questionnaire item, respondents specify their level of agreement to a statement. The scale is named after Rensis Likert who published a report describing its use.[17]

Measurements of central tendency: mean, mode and median, see.Table 2.

Measurements of dispersion: see. Note that variance is an example of dispersion.Table 3.

Nominal data: is the least robust data which categorises, but does not hierarchically rank data into mutually exclusive categories.

Normal distribution: when a large number of measurements are made at random of one particular variable, the results usually fall into a pattern. Most of the measurements will lie close to the mean value, with few values lying at the extremes. When a frequency distribution is plotted a familiar bell shaped curve is produced which represents a normal or Gaussian distribution.

Ordinal data: is where the data has a clear order or hierarchy but not on a calibrated scale. The ordinal data can be ranked (such as from tallest to smallest) or ordered (e.g. data categorised via the level of agreement with statements on a Likert scale e.g. very painful, painful etc.). Both numerical and non-numerical can constitute ordinal data.

Parameter: a ''true'' measurable characteristic of the population that cannot, in practice, be known with certainty, e.g. the mean breast thickness of all women eligible for mammography breast screening. A ''statistic'' i.e. the mean value in a sample of patients is measured and used as an estimate of the true mean of the population.

Parametric tests: are selected for data that is normally distributed;

Pie charts: used to demonstrate distribution ratios as it is easy to compare the relative sizes of the various components.

Table 3 Descriptive statistics for each level of measurement.

Level of measurement	Description	Measure of central tendency	Measure of dispersion
Nominal	Classification	Mode	Frequency distribution
Ordinal	Relative rankings	Median Mode	Frequency distribution Percentile, Maximum and minimum, Range
Continuous	Rank ordering with equal intervals	Mean[a] Median Mode	Frequency distribution Percentile Maximum and minimum, Range, inter-quartile range, Standard deviation

[a] If the distribution of data is not normal, e.g. there are a number of outliers that influence the mean dramatically (making the data skewed), it is sometimes better to present the median to give a better reflection of the value of most of the data points.

Population: a complete set of individuals, objects or measurements having some observable characteristic in common e.g. all women who have had one child.

Ratio data: is the strongest data, and has a true zero value. It can be achieved by the use of a calibrated scale to provide quantitative measurements.

Sample: a subset of the population, selected to participate in a study.

Standard deviation: figure that gives an indication of how closely distributed measurements or observations are around the mean of all the measurements. It gives an indication of the variance of the data. A small standard deviation will result in a steep-curved bell curve when presented in a graph. Oppositely, a large standard deviation will result in a flat bell curve.