The role of statistics in wound care

John Stephenson
Reader in Biomedical Statistics, School of Human and Health Sciences, University of Huddersfield, Huddersfield, UK

Key words
- Statistics
- Study design
- RCT
- Analysis

Wound care, like virtually all branches of medical care, is a multidisciplinary subject. Over the last few decades, there has been an increase in the number of clinical academics who have a role within healthcare and academia. Hence, a typical multidisciplinary team will now incorporate non-clinicians that include, for example, a biomedical statistician, health economist and data manager, alongside clinical staff. While the non-clinicians are important members of the team, all too often, neither the clinical nor non-clinical staff have much of an idea of the role played by each other. As a (non-clinical) biomedical statistician myself, in this article, the importance of the role of statistics and the statistician in wound care studies will be briefly summarised.

Study design

Biomedical statisticians are instrumental in designing robust clinical trials and observational studies in wound care. In fact, many biomedical statisticians prefer to be involved in the design process from the outset rather than be given data to analyse that have been already collected, as such data may have been collected with insufficient consideration of its application to a particular research question. The biomedical statistician can help decide on the study design: whether the study should be, for example, a randomised controlled trial (RCT), cohort study, case-control study or some other type of study design, depending on the research question. For example, let's say we are planning to assess the efficacy of hydrocolloid dressings (against usual care, such as a standard gauze dressing) in the healing of diabetic foot ulcers, this research question is most appropriately addressed using an RCT, with participants allocated to either a hydrocolloid dressing or a standard dressing group.

Another example may be a study to assess the effect of the provision of patient information on self-management of chronic wounds, measuring patient confidence in self-management before and after receiving the information. This research question is most appropriately addressed using a different study design altogether: a pre-post study in which patients act as their own controls. The early identification of the study design is necessary as the statistical processes and tests that will be followed when data have been collected depend largely on the design chosen.

Sample size calculation

The first actual use of statistics in a wound care project is normally in the sample size calculation required by many studies. The aim here is to get the right number of patients for our study. Too many, and time is wasted (our time, and our patients' time) and resources in recruitment and follow-up of patients who we did not need; too few, and we may be conducting an underpowered study: one which may fail to detect a real treatment effect, leading to an incorrect inference that a potentially beneficial wound care treatment is ineffective. As any study may expose patients to potential risks, such as infections, pain, or adverse reactions, if a study is too small to produce meaningful results, participants may face risks without any chance of contributing to medical knowledge.

Getting the sample size right is difficult. Let's say we want to estimate the sample size needed in an RCT, we will need to estimate the anticipated treatment effect (treatments with large effects need fewer patients to show their worth than treatments with small effects, as there is less chance of a large effect being lost in any random noise in the data). This, in turn, requires some knowledge of the expected outcomes in both study groups (usual care and new treatment), and the variability of the data we will collect. We also need to state the significance level of the testing process and the power we wish to achieve. In the context of clinical science, "power" means the probability that we will detect any real treatment effect that may exist and not mistake it for random noise in the data. Obviously, we want this probability to be as large as possible, but it comes at a cost: a higher probability of detection means more patients. While 100% power is not feasible, 80% power is generally achievable and is a fairly standard requirement.

More elaborate study designs need yet more parameters feeding into the sample size calculation. For example, many wound care studies involve data collected from multiple sites. It may not be feasible for a particular site to offer both usual care and a new treatment

Submission date
14 February 2025
Accepted date
8 April 2025
Published date
29 April 2025

Declarations
The author has no conflicts of interest to disclose.

Corresponding author
John Stephenson, University of Huddersfield, Queensgate, Huddersfield, UK, HD1 3DH. Email: J.Stephenson@hud.ac.uk

under assessment, so randomising sites (rather than individual patients) to either usual care or a new treatment is commonly done. Using this method, all patients at a particular site receive the same treatment. This has pragmatic advantages but complicates the sample size calculation process, as we also now need to factor in a clustering effect, in which we have to estimate any degree of self-similarity between patients within each site.

Sample size calculations is an area where the biomedical statistician will normally need to liaise with clinical colleagues on the research team, who will generally have a much better idea of anticipated effects. However, even the clinicians cannot read the future, and the sample size calculation is essentially about estimating parameters in studies that have not yet happened. Sometimes a small pilot study can be helpful to arrive at least ballpark figures for these estimates.

Randomisation

If the study being proposed involves randomisation of treatments, the next use of statistics is in the randomisation process. The allocation of patients to treatment is usually conducted by the statistician, who will not normally be involved directly with patients: if the randomisation were done manually by investigators, they might unintentionally influence assignments based on patient characteristics. For example, in a wound care trial, a member of the trial staff might be tempted to assign patients with larger ulcers to the treatment group, skewing results.

Although patients can be allocated to treatments by, for example, flipping a coin, it is rarely quite as simple as that: coins are not guaranteed to land heads 50% of the time! A study of, say, 30 patients, in which 23 receive usual care and only seven get the new treatment, just by the way the coin falls, is unlikely to be very informative. Statisticians generate concealed allocation sequences that researchers and participants do not see. This helps maintain blinding, preventing investigator bias in treatment administration and outcome assessment. A study can often be strengthened if the statistician themselves are blinded to the group identities: they may receive data from the trial coordinator by groups denoted simply Group A and Group B. The statistician is then unable, unwittingly or otherwise, to use their knowledge of group allocation to influence the result of the trial.

Preliminary statistical procedures

In common with other clinically based research studies, many wound care studies are not able to collect all the data intended. Patients may drop out of a study early, move away or die. Staff may be off sick, or equipment may be faulty. Either staff, patients or both may deliberately withhold information they consider to be sensitive.

While it is tempting to ignore missing data, some types of missing data can lead to problems. Maybe some patients stopped using a wound dressing because they found it uncomfortable, so the only ones remaining to be analysed are those who did not have problems. If we ignore the missing values, we might overestimate the comfort levels associated with that dressing. Another example might be a clinical trial testing the effectiveness of a new wound dressing for diabetic foot ulcers, in which patients are scheduled for follow-ups every 2 weeks to assess healing progress. However, some patients with severe, slow-healing wounds stop attending follow-ups because they are discouraged by the lack of improvement. Again, only those with positive outcomes remain to be analysed, potentially leading to biased conclusions.

Many statistical strategies can be utilised here to ensure that, as far as possible, inferences are not biased by the missing data. For small amounts of data missing completely at random, a statistician may choose to conduct a complete case analysis (considering only participants who provide a full set of readings). In other cases, non-essential variables may be deleted, or missing data may be imputed, normally involving a computational algorithm. Larger amounts of missing data and, in particular, data that are not missing at random can be problematic to deal with.

Inferences from many clinical studies go awry due to failure to detect data errors. Wound care data are often collected by time-pressured nurses. It is all too easy to write down a wound length, as, say, 50 cm rather than 50 mm; mistype a patient's age as 95 years rather than 59 years; put a tick mark in a questionnaire between two boxes rather than inside one of them; or write an entry on a patient record form with a 6 that looks like a 0. Equipment can also be faulty or mis-calibrated. However, most of these errors are not easy to spot at a glance. Nor is it always easy to distinguish between an error arising from a data entry or transcription error and an unusual but genuine result, which should not be amended or removed. Close monitoring of a study is essential to allow detection and distinction of such cases. We can apply statistical methods to check the provenance of our data and make sure before we conduct the

main analysis, that the data set we are working on is of good quality.

Data analysis

The main data analysis is normally completed after the final patient has completed their last follow-up appointment. While study statisticians are not normally directly involved in data collection, they have a role to play in advising clinical colleagues who are involved in this process to ensure that the data collected are suitable for subsequent analysis.

Many, if not most, wound care studies involve the collection of quantitative data. This comprises variables that can be either numerical or categorical. A typical numerical variable might be, for example, a wound length or width or a time-based measure, such as the time to 50% reduction in wound size. A typical categorical variable might be, for example, the status of the wound (healed or unhealed) within 30 days of a baseline assessment, the type of a wound (e.g. categorised as a diabetic foot ulcer, leg ulcer, pressure ulcer) or treatment status (new treatment or usual care). Variables can be related to the wound itself (diameter, anatomical location etc.) or the patient with the wound (e.g. age, sex, ethnicity). They can be "outcome" variables (also variously known as endpoints, criterion variables, response variables and dependent variables), such as pain, wound diameter or quality of life, or "predictors", i.e. those variables that may predict or explain an outcome (also variously known as explanatory variables, covariates, factors or independent variables), such as type of treatment received or patient comorbidities. They can be self-reported by the patient or clinician via questionnaires or surveys (pain, type of dressing used etc.) or collected via medical devices or other equipment (negative pressure, systolic blood pressure, body mass index etc.).

All these variables require some kind of statistical treatment! In fact, almost the only variables collected in wound care that do not require the use of statistics are responses to open-ended questions on questionnaires eliciting, for example, patients' opinions, feelings and values, for which we need the assistance of our qualitative colleagues.

Almost all clinical studies include a descriptive summary of data. In a wound care context, this usually means both participant demographics and wound parameters, as recorded at baseline and at one or more subsequent time points. Many studies also require some form of inferential statistical testing, usually, if the intention is to generalise findings from a sample (the patients who we actually access in our study) to a population (the much wider body of individuals to whom we believe our findings will apply). Different study designs require different procedures, although the basic aim in most cases is the same: to assess the effect of interest in various ways. An "effect" could be an observed difference between study groups (such as signs of localised infection in a control group and a group where patients are treated with an antimicrobial dressing), the difference between a measure taken at baseline and post-intervention from a single patient group (such as pain levels during and after surgery), an observed relationship between two variables (such as the extent of mobile health technology use and wound care knowledge) or many other quantities.

A recent example from my own work is an RCT conducted to assess the effectiveness of a heel boot protector in reducing hospital-acquired heel pressure ulceration in 394 intensive care patients in three hospitals in New South Wales, Australia (Barakat-Johnson et al, 2022). In this study, we compared outcomes in an intervention group (patients fitted with the heel boot) and outcomes in the control group, comprising patients with heel offloading using pillows, which was the standard mode of care. Both the sample size calculation and the subsequent statistical analysis were made slightly more problematic by the study design: we based our analysis on patient heels rather than whole patients — and, of course, each patient contributed two heels to the study (except any who had pre-existing pressure ulcers on one or both heels). This results in an effect similar to that found in some multisite studies — data are clustered within patients.

We used as our primary outcome hospital-acquired heel pressure injury within 28 days from intensive care unit admission. We also measured several secondary outcomes. The statistics relating to the primary outcome at the end of the trial looked initially impressive, with 11 new injuries in the control group (197 patients) and just one in the heel boot group (also 197 patients). But the statistics did not stop there: we needed to verify that the hazard of pressure injury incidence was significantly lower in the heel boot group, i.e. we needed to be confident that the effect we had seen was not likely to be a chance finding, so we could infer that what we had found was a true reflection of the worth of the intervention. This inference is what statisticians mean when they talk about "significance": they do not simply mean "important" or "large in magnitude"; they are referring to a situation where the result of a study or experiment is judged to be sufficiently

unlikely to have happened by chance for the "chance finding" explanation to be accepted. A significant result means that the evidence suggests an effect is real rather than random. Significance is quantified with a *P*-value, which crops up in many clinical research papers: the lower the p-value, the stronger the evidence for a real effect, with *P*-values below 5% usually taken to imply significant results.

Fortunately, the effect measured in our heel boot study of 10 fewer cases in the heel boot group could indeed be shown to be significant, with a calculated *P*-value of 2.39%, comfortably below the 5% (0.05) threshold. This establishment of statistical significance required a test of a hypothesis. We usually test a null hypothesis (of no effect); for example, that the difference in means in two population subgroups is zero. It is clear from the data, but important to note nonetheless, that the findings represent a positive benefit of the heel boot: we have a positive direction of association.

Statistics is also useful for estimating the precision of an effect. As we are usually constrained by time and resources to analyse only a very small proportion of the population of interest, we will never know what the true effect on the whole population might have been. In our study, the population might be every single patient admitted to an ICU in any hospital in New South Wales, or maybe anywhere in Australia, or maybe even anywhere in the whole world. The numbers involved make clear that our sample of 394 patients may seem large at first glance, but it is actually a negligible fraction of the patient population it represents. While we can certainly use the findings from our sample (hazard of heel pressure ulceration in ICU patients with heel boots about 9.0% of the corresponding hazard of heel pressure ulceration in ICU patients with pillow offloading) as the best estimate of the state of affairs in the population, it would be unwise to assume that this precise figure would apply across an ICU patient population of possibly many hundreds of thousands.

Instead, a further bit of statistical analysis is used to calculate what is known as a confidence interval: a range within which we have a certain degree of confidence that the true value in the population would be found, were we able to measure it. Our study found that this interval was actually quite wide: in the underlying population, the hazard of pressure injury when fitted with a heeled boot is, we believe, with a certain degree of confidence, somewhere between 1.1% and 72.7%. But whatever the true effect of the heel boot is, we are confident that the hazard of pressure ulceration of patients wearing it is less than 100% of the hazard of patients using pillow offloading! Thus the heel boot represents a significant and substantive improvement in terms of the reduction of associated pressure injury incidence.

So those are the "big 4" statistics you would want to see in any comparative wound care study: the magnitude of the effect, the direction of the effect, the significance of the effect and the precision of the effect. Although this study was an RCT, I would be looking for the same four pieces of information in many other comparative study designs, such as a cohort study, case-control study or pre-post study. In other studies, notably ungrouped (non-comparative) studies, I may be looking for different pieces of information: for example, in a study investigating the relationship between blood glucose levels and wound healing time, I might be looking for a correlation coefficient (telling me the strength of the relationship between these variables), the significance of that relationship, an estimate of the effect of a unit change in blood glucose levels on healing time, and a predicted estimate of the healing time in a patient with a give blood glucose level. Other, more complex studies, such as those with multiple treatment groups, imprecisely known outcomes (such as when an outcome is the time to an adverse event in patients who are monitored infrequently), a or series of observations made on the same patients, might lead to further different sets of statistics to be generated.

Every wound care study is different, and the biomedical statistician needs to choose both the statistical procedure and the way it was to be represented carefully: different statistical tests and different methods of graphical presentation and tabulated measures are valid in different types of studies. However, most procedures can be easily implemented using standard statistical software.

Although data analysis after completion of all patient appointments is the norm, data analysis does not necessarily have to wait until a study is completed. Many wound care studies utilise interim analyses conducted while the study is still in progress. Many trials can and should be stopped early for clear indications of either the efficacy or futility of the treatment being tested. However, the extent to which a new treatment can be confidently stated to be either effective or ineffective before all data have been collected is not easy to establish. Timely statistical intervention can stop a trial of an obviously effective or ineffective treatment, thus preventing further patients from entering a study unnecessarily.

Conclusion

The conduct and interpretation of statistical

procedures are essential to the analysis of wound care data and allow us to answer quantitative research questions. These inferences can target interventions to improve patient outcomes. For example, if statistical analysis indicates that a particular wound treatment is more effective in a specific patient demographic, healthcare providers can tailor their approaches accordingly. Additionally, statisticians contribute to cost-effectiveness analyses, helping policymakers allocate resources efficiently.

Statistics and biomedical statisticians play a fundamental role in advancing wound care by designing rigorous studies, analysing complex data and evaluating treatment efficacy. Judicious use of statistics in wound care studies ensures that clinical decisions are backed by robust evidence, ultimately improving patient outcomes and shaping the future of wound management. As wound care continues to evolve with technological advancements and data-driven innovations, the role of statisticians will remain indispensable in driving progress and enhancing healthcare delivery. Through collaboration with healthcare professionals and policymakers, statisticians will continue to contribute significantly to the optimisation of wound care practices and patient wellbeing.

Reference

Barakat-Johnson M, Lai M, Stephenson J et al (2022) Efficacy of a heel offloading boot in reducing heel pressure injuries in patients in Australian intensive care units: a single-blinded randomised controlled trial. *Intensive Crit Care Nurs* 70: 103205. https://www.sciencedirect.com/science/article/abs/pii/S0964339722000106?via%3Dihub