As is true for other clinicians, research skills for anaesthesiologists are important not only for individual progress, but also for advances in this fast-evolving discipline. This is easier said than done. Anaesthesiologists are the busiest among clinicians, first among equals as it were. I should know having worked as an anaesthesiologist after graduation, first as a resident and later as ad hoc anaesthesiologist in the armed forces before switching over to public health and epidemiology.

After a busy day in the operation theatre (OT), even if the spirit is willing, the flesh is likely to be weak and tired. To add to the hurdle, research methods, particularly biostatistics is like a "black box" to most clinicians, including anaesthesiologists. Cutting the clutter can overcome the odds to a large extent.

First, what is research? In my introductory sessions in research workshops this is my first question. Few could answer clearly. Simply put, there are two steps in the research process, namely, observation and measurement. My subsequent questions are “Has anyone ever carried out research and if so, when did they first start?” Few hesitantly mention postgraduate dissertation or some student project, again short of satisfactory answers. The simple fact that research starts from birth eludes most. Every child is born a researcher as observation, the first step in research, is keenest in the child.

The little girl may observe her mother enter the room and press the light switch, which illuminates the whole room. A single observation may stimulate her interest. Subsequent observations may suggest a pattern of ‘cause-effect’ relationship, the switch being the cause followed by illumination of the room, the effect. The child has performed its first observational study!

To confirm the ‘cause-effect’ relationship, the child may press the switch when no one is watching and get ecstatic when the room gets illuminated. The child has performed its first interventional study! One can imagine her confusion when occasionally pressing the switch leads to nothing because of a power cut. The child has encountered her first confounder or effect modifier!

Research in later life is based on these traits of human curiosity, creativity, and exploration all of us are born with, only to atrophy with formal education that imparts knowledge but often lags in preserving creative thinking. History of research has examples of innovators who dropped out of formal education and preserved these traits leading to discoveries bringing about a paradigm shift.[1]

For rigor in research, the innate observation skills have to be refined. In clinical research, observations are made on a sample of patients. This can lead to ‘selection bias’ if the sample is not representative of
the target population that is, the people to whom the study results may be generalised. This selection or sampling bias can limit the extrapolation of the study findings or its external validity. The results may hold true only for the type of patients in the study sample who may differ from patients seen in the community. So, while planning a research or appraising one, we should take a hard look at how the sample was selected and whether it represents the target population. It is a matter of judgment rather than statistics to decide external validity that is, the generalisation of the findings.

For research to be replicable, the next step of research, that is measurement is equally important. These should be described in an easy to understand and logical way. The methods of measurements of all variables, that is, predictor, outcome and confounders should be described in sufficient detail which can enable others to replicate the study. The importance of attention to details in this aspect is most important, more than any complicated statistics applied to the data later. Proper measurements are the foundation of the research project. In addition to imprecise measures, defective reagents, instruments, measurement biases can also creep in due to human errors both among participants (placebo effect) and the investigators (cognitive biases, confirmation bias, inter and intra-observer errors). If measurements are compromised, the study cannot be salvaged by statistical techniques. Incorrect statistical techniques can be redone but incorrect measurements cannot. While a representative sample gives external validity to the study, good measurements assure internal validity, that is, the data conveys the truth in the sample. Without internal validity, the study cannot have external validity either.

The third bone of contention in research is confounding variables. These variables can distort the inference of cause-effect relationship and give rise to spurious associations. For instance, in a study of efficacy of a vaccine, if we compare the disease incidence between those vaccinated and those not vaccinated, we may on finding that the incidence of disease is less in the vaccinated conclude that the vaccine is effective. However, if we do not match the socioeconomic status, this may well be due to this confounding variable because people with higher socioeconomic status are more likely to be vaccinated and also more likely to have better access to good housing, transport and quality of life, which may reduce the disease incidence in them rather than the vaccine.

Known confounders like socioeconomic status, gender, age, literacy status can be matched while planning the study. Failing which, they can be corrected for by statistical techniques such as stratified or multivariate analysis later. More problematic are unknown confounders. The only way to eliminate the distorting effects of unknown confounders is a randomised controlled trial (RCT) which has been called the gold standard of evidence-based medicine.

The most intimidating but perhaps of lesser importance is statistics in clinical research. If earlier biases such as selection, measurement and confounding are taken care of while collecting data, one can explore any number of statistical techniques without irreversible harm. Data can be re-analysed.

Statistics start with sample size calculation. The clinical researcher has to give certain inputs to the statistician or the statistical software which are frequently used nowadays. These are Type 1 error (or alpha), Type 2 error (or beta), power of the study (which is 1 – Type 2 error) and effect size. Type 1 error is the probability of getting a difference when in truth there is no difference between the two groups. Type 2 error is the opposite that is, failing to detect a difference when actually there is a difference. By convention Type 1 error is kept at 5% (P = 0.05), and Type 2 at 20% (power = 80%). The sample size also depends on the effect size that is, the minimum difference in the two groups which would be clinically relevant. Smaller the difference, the larger will be the sample size to detect it. The effect size of clinical relevance can only be decided by the researcher and not by the statistician or the software.

The next statistical part is the analysis of the data. This can be divided into descriptive statistics and inferential statistics. Methods of describing both types of statistics depend on the type of data. Data can be counted (male/female, cured/not cured, dead/alive), ranked (grades of muscle power, American Society of Anesthesiologists (ASA) physical status classification system), or measured (pulse rate, blood pressure, weight). They are called categorical, ordinal, and interval data, respectively. For categorical data, descriptive statistics are proportion or percentages with 95% confidence intervals. Ordinal data are summed up by median and range, while the most
robust interval data are summarised with mean and standard deviation (SD), along with 95% confidence intervals.

The choice of statistical tests for hypothesis testing,\(^7\) also depends on the type of data. For categorical data, the Chi Square test gives the \(P\) value. However more importantly the Relative Risk (RR) for cohort studies and Odds Ratio (OR) for case-control and cross-sectional studies with 95% confidence intervals are more desirable. For ordinal data, the statistical significance is tested by non-parametric tests such as Kruskal Wallis or Mann-Whitney depending on the number of groups. The most robust tests of significance are the parametric tests such as Analysis of Variance (ANOVA) (for more than two groups) and the T-tests for two groups which can be applied only on interval data.

The most important point to be noted is that all estimates on the study sample are point estimates. When reporting the results, the point estimate should be stated along with the 95% confidence intervals. Proper use and interpretation of confidence intervals are more important than just the \(P\) values which can be dispensed with if the confidence intervals are stated. For example, when the 95% confidence interval for OR and RR includes ‘1’ we can infer that there is no statistical significance in the difference. Similarly, when the 95% confidence of the effect size in a parametric test includes ‘0’, the results are not significant. Increasingly, journals are insisting on presenting statistical results with the 95% confidence intervals and rightly so. The confidence intervals convey much more than just the \(P\) value. They indicate the adequacy of the sample size. Too wide confidence intervals indicate lack of precision of the estimates as a consequence of small sample size. Therefore, all researchers and reviewers should be conversant with the concept of confidence intervals.\(^8\)

The last few paragraphs may have been difficult for readers not very conversant with statistical methods. They may find it reassuring to know they contain most of the statistical techniques needed to develop a ‘statistical sense’ enabling a fruitful collaboration with biostatisticians and also optimum use of statistical software. Going through the papers listed in the references will further give clarity to these concepts.

REFERENCES

1. Banerjee A. Do research workshops lead to research: Or is it like lecturing birds how to fly? Med J DY Patil Vidyapeeth 2018;11:383-4.
2. Hulley SB, Newman TB, Cummings SR. Getting started: The anatomy and physiology of clinical research. In: Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB, editors. Designing Clinical Research. 4th ed. Philadelphia: Wolters Kluwer, Lippincott Williams & Wilkins; 2013. p. 2-13.
3. Browner WS. Methods. In: Publishing and Presenting Clinical Research. 3rd ed. New Delhi: Wolters Kluwer, Lippincott Williams & Wilkins; 2012. p. 27-43.
4. Newman TB, Browner WS, Hulley SB. Enhancing the causal inference in observational studies. In: Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB, editors. Designing Clinical Research. 4th ed. Philadelphia: Wolters Kluwer, Lippincott Williams & Wilkins; 2013. p. 117-36.
5. Hariton E, Locascio JJ. Randomized controlled trials – The gold standard for effective research. BJOG 2018;125:1716.
6. Noordzij M, Tripepi G, Dekker FW, Zoccali C, Tanck MW, Jager KJ. Sample size calculations: Basic principles and common pitfalls. Nephrol Dial Transplant 2010;25:1388-93.
7. Banerjee A, Chitnis UB, Jadhav SL, Bhawalkar JS, Choudhury S. Hypothesis testing, Type I and Type II errors. Ind Psychiatry J 2009;18:127-31.
8. Schober P, Bossers SM, Schwarte LA. Statistical significance versus clinical importance of observed effect sizes: What do \(P\) values and confidence intervals really represent? Anesth Analg 2018;126:1068-72.

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

How to cite this article: Banerjee A. Research and the anaesthesiologist: Cutting the clutter and overcoming the odds. Indian J Anaesth 2021;65:183-5.