National Medicines Information Centre

VOLUME 28
NUMBER 1
2022

St James's Hospital, Dublin 8
Tel 01 4730589 or 0818727727; twitter @NationalNmic; nmic@stjames.ie

HOW TO REVIEW A CLINICAL PAPER

- **Evidence-based medicine involves the use of current best evidence in the clinical management of individual patients**
- **Clinical guidelines incorporate best available evidence and are developed to inform both the physician and patient about appropriate healthcare choices**
- **Statistical significance does not always equate to clinical relevance**
- **A structured approach is advised when reading clinical research papers and guidelines to assess the clinical applicability of the evidence for individual patients**

INTRODUCTION

Evidence-based medicine (EBM) may be defined as the "conscientious, explicit and judicious use of current best evidence in making clinical decisions about the care of individual patients".[1] The focus of EBM includes critical appraisal of the literature, the development of systematic reviews and clinical practice guidelines.[2] Ideally, clinical decisions should be based on the totality of current best evidence, gathered from the results of individual clinical trials or studies.[2,3]

Clinical effectiveness is a collection of activities and tools (e.g. guidelines and audit), based on research and measurement, that are used to improve the quality of healthcare and improve patient safety.[4] In Ireland, clinical effectiveness guidelines are developed through the National Clinical Effectiveness Committee (NCEC) that was set up by the Minister for Health in 2010;[4] these clinical guidelines are available at https://www.gov.ie/en/collection/c9fa9a-national-clinical-guidelines/.[5] **Clinical guidelines** are systematically developed statements, based on a thorough evaluation of the available clinical evidence, to inform both the physician and patient about appropriate healthcare choices for specific individual clinical circumstances.[4] The implementation of clinical guidelines has been shown to improve health outcomes for patients.[4]

An understanding of the processes used to evaluate the clinical evidence is important to enable healthcare professionals to determine the relevance of any clinical guidance for their own individual patients.

This bulletin (which updates an earlier NMIC bulletin[6]) outlines the basic elements of clinical research and identifies the key points in the critical review ("critical appraisal") of published clinical research papers.

ASSESS THE STUDY TYPE

Clinical research may be subdivided into **primary** (which involves "original" or new clinical studies) or **secondary** (which involves the use of existing studies i.e. a review of a number of primary studies fulfilling certain eligibility criteria).[7] Most of the published clinical research is "**quantitative**" which describes studies that collect numerical data, which is the focus of this bulletin.[8] There are two broad types of quantitative study in clinical research: experimental (or interventional) and observational (or epidemiological) - see table 1.[9] The study type and design are determined by the clinical question under investigation.[3]

Table 1: Types of primary clinical research[9]

IS THE INTERVENTION ASSIGNED?	
YES ↓	**NO** ↓
Experimental / Interventional study ↓ Randomised controlled trial (RCT) Non-randomised /(open) controlled trial Single arm trial	**Observational study** ↓ Cohort study Case-control study Cross-sectional study Case reports / case series

EXPERIMENTAL RESEARCH

Experimental research involves the researcher intervening in some way and evaluating the outcome of that intervention. The most definitive evidence for the safety and effectiveness of a therapeutic intervention is provided by a **randomised controlled trial** (RCT), which is recognised as the gold standard in experimental research, especially for treatment regimens.[7,9] In a therapeutic RCT, participants are **randomly assigned** to one of two or more treatment arms (i.e. random allocation means the allocation to any arm of the trial is independent of selection by the participant and investigator); participants in each arm are then managed (**controlled**) in an identical manner with the exception of the treatment received, and their responses are compared.[7,8] This removes the potential for selection bias by the investigator and means that any difference between the study arms identified in the trial is likely to represent a true difference between the treatments.[9] Many RCTs also involve masking ("blinding") of the assignment of subjects; in a **double-blind study** neither the subject nor the investigator are aware of the treatment allocation, which reduces the risk of evaluation bias.[10,11] The control group in a RCT may be allocated to no active treatment (a **placebo-controlled trial**) or to an existing treatment (an **active comparator trial**).[8] Placebo-controlled trials help to establish the true efficacy of a treatment regimen while active comparator trials help to determine the efficacy of the test treatment, relative to an existing established treatment. In clinical research it may not always be ethically possible to undertake a placebo-controlled trial (e.g. in the management of serious conditions for which an existing therapy exists and which should not be withheld); neither may it always be possible

to undertake a randomised trial.[7] **Non-randomised clinical trials** may be subject to selection and/or evaluation bias and therefore the information they provide is regarded as supplementary to RCTs.[7]

OBSERVATIONAL RESEARCH

Observational research involves the researcher observing certain aspects of an existing situation (e.g. a therapeutic regimen) without intervention; the investigator has no role in assigning therapy.[12] The **main observational study types** are as follows:

Cohort study: Subjects are followed up over time (often many years) to observe the effect of an exposure (e.g. to tobacco), the natural history of ageing, or of disease aetiology in a specific cohort of subjects.[9,13] Examples include the national longitudinal study of children, "Growing up in Ireland", and TILDA (The Irish LongituDinal Study on Ageing).[14,15]

Case-control study: This type of study works backward from an outcome (e.g. disease) in order to identify a *possible association* between the outcome and a particular exposure (e.g. medicine). Patients with the disease are matched (usually by age and gender) with a group of "controls" without the disease, and the level of exposure to the suspected agent is identified for each group of subjects.[9,12,16] Case-control studies are useful in assessing disease aetiology; they have been used to evaluate potential drug safety issues e.g. risk of venous thromboembolism (VTE) with use of combined hormonal contraceptives (CHC).[17]

Cross-sectional studies are descriptive studies which provide a snapshot in time of the issue under investigation.[9,18] Examples include screening for the prevalence of cardiovascular (CV) disease / risk factors (e.g. EUROASPIRE IV study)[19] and lifestyle and attitudes surveys (e.g. the Healthy Ireland Survey 2021, providing an insight into the impact of COVID-19 restrictions on the health and wellbeing of the people of Ireland).[20]

Case reports / series are uncontrolled observations of a single subject / group of subjects with a shared condition; these may be useful in identifying early potential drug safety signals.[9,18] Examples include the reports of teratogenicity with use of thalidomide, and hepatotoxicity with use of nimesulide.[21,22]

Limitations: Observational studies are susceptible to **potential biases** which can be defined as any tendency to distort or affect the results of a study other than the exposure.[12] These include: (1) **selection bias** (e.g. in a cohort study, are the participants in the exposed and unexposed groups similar in all important respects apart from the exposure?, and in a case-control study, are the cases and controls similar in all important respects except for the disease in question?); (2) **loss to follow-up / non-responder bias** (any potential differences between responders and non-responders cannot be captured); (3) **information bias** (e.g. in a cohort study, whether the information about the outcome is obtained in the same way for those exposed and not exposed, and in a case-control study, whether information about exposure is gathered in the same way for cases and controls) and (4) **recall bias** (e.g. differences in recollection of exposures among the cases compared to controls).[9,12,23]

Confounding, another type of bias may be defined as any factor which distorts the association between the exposure and outcome,[12,23] and may also alter the validity of results (e.g. a high BMI and VTE risk in women taking CHC).[12,21] In the case of drug safety issues, it may be **difficult to confirm a causal association** between use of a medicine and a noxious outcome, based on observational studies.[12] Therefore, healthcare professionals should give consideration to all sources of potential bias and their impact on the results when reviewing observational research.

SECONDARY RESEARCH

Secondary research focuses on reviewing primary research.[24] It involves defining a specific research question and performing a **systematic review** of a group of primary research studies, in order to look for consistency in the findings from the individual studies.[25,26] Systematic reviews can summarise not only RCTs, but also cohort studies, case-control studies and even case reports.[2] A **meta-analysis** involves the pooling of results from these primary studies using statistical analyses; it provides an overview of the results, with measures of uncertainty around these results using 95% confidence intervals (CI), with regards to a particular outcome measure.[25,26] A systematic review of all available evidence is always more reliable than any single study, provided the review is properly conducted. Systematic reviews are essential for developing clinical practice guidelines, for avoiding duplication of research efforts and for helping inform design of new research studies.[2] Table 2 outlines the essential elements of a systematic review.

Table 2: Steps involved in a systematic review[24,25,27]

- Formulate clear objective(s) of the review (e.g. PICO* format)
- Undertake a search of the literature, using clearly defined criteria (e.g. types of studies to be included, sources to be searched)
- Apply clearly defined inclusion / exclusion criteria for all identified studies
- Seek additional information from primary researchers if possible (especially important for meta-analysis)
- Undertake the review of each study using explicit criteria (to evaluate the quality [e.g. risk of bias] of research)
- Analyse the pooled data using validated methods: either systematic critical appraisal or meta-analysis
- Publish the findings, using a similar format to that used for primary research papers

*PICO=**P**opulation of interest; **I**ntervention; **C**omparator; **O**utcome

The Cochrane Collaboration is an independent international organisation, consisting of global experts, which undertakes systematic reviews of the effects of various healthcare interventions.[28] These reviews which are regarded as one of the most enduring and reliable systematic reviews,[24] are valuable sources of information for decision makers and researchers, as well as patients.[28] **Access to the online Cochrane Library of systematic reviews is available free of charge in Ireland**, via **www.hrb.ie**.

PRACTICAL CONSIDERATIONS WHEN REVIEWING CLINICAL PAPERS

Most clinical papers are presented in a standard [IMRaD] format: Introduction (why the study was done), Methodology (how the study was done and what statistical analysis was used), Results (what was found) and Discussion (what the results mean).[7]

The **introduction section** of a paper should provide the background to the research and highlight the aim(s) of the study, including details on the **P**opulation of interest; **I**ntervention carried out; **C**omparator used (if appropriate); **O**utcome(s) evaluated (**PICO**).[6,29]

STUDY DESIGN ASPECTS

The **study design** is determined by the question being investigated. Table 3 summarises the preferred study

design appropriate to the research question under evaluation.

Table 3: Preferred study design according to the research question[7,29]

Research question	Preferred study design
Therapy or any intervention e.g. testing the efficacy of drug treatments, surgical treatments	Randomised controlled trial / systematic review
Diagnosis e.g. measurement of a condition and determining if a diagnostic test is valid and reliable	Cross-sectional study
Prognosis e.g. determining how to predict a patient's clinical course following exposure to an agent	Cohort study
Association/Causation e.g. determining whether an agent may be associated with a disease	Case-control study / case series

Normally the aim of a comparative clinical trial is to show a difference between arms of the trial or to "reject" the **null hypothesis** of no difference. The study sample size is calculated statistically to achieve this aim.[7,10] This calculation takes into account several aspects including the study design, and the level of departure from the null hypothesis (i.e. **Minimal Clinically Important Difference** (MCID) in effect between treatments).[30] It is important to check that the MCID used in a trial is clinically relevant for the condition under evaluation (e.g. blood pressure drop of 10mmHg between antihypertensive agents) and supported by evidence (from literature or pilot studies), since the **power** of a study to find a difference if one truly exists relates to the chosen MCID as well as sample size.[31] **Statistical significance** does not always equate to **clinical relevance.**[11]

Other types of RCTs can be designed to show: (1) **superiority** (i.e. a significant difference in one treatment arm over another, in terms of the chosen MCID), (2) **equivalence** between treatment arms (i.e. "no better or no worse", in terms of the equivalence margin; clinically this margin is not considered an important difference) or (3) to show that any difference is not clinically inferior, in terms of the MCID (**non-inferiority** trial).[8,32,33] Many of the pivotal studies in drug development use a non-inferiority design; examples include the clinical trials which formed the basis of approval for the direct oral anticoagulant agents (e.g. apixaban compared to warfarin).[34-36] Finally, an important aspect of the study design is the **choice of study population**, since this will determine the applicability of the study results for the larger population of patients. This is defined as the study inclusion and exclusion criteria.

HOW TO INTERPRET CLINICAL STATISTICS

The **methodology** section should contain detailed information about how the study was conducted (how all subjects were managed, evaluated and followed up) and how the results were analysed. Table 4 summarises statistical terms frequently encountered when reading a clinical paper. Justification for the sample size calculation should usually be presented in the methodology section of a clinical paper. Most clinical trial findings should be analysed on the basis of the **intention-to-treat** (ITT) principle, where study subjects are analysed as members of the trial arm to which they were randomised, irrespective of whether they received or adhered to that arm (e.g. treatment regimen) during the study.[8] They may also be analysed according to what they actually received (**per protocol analysis** [PPA]).[8] ITT is said to reflect "real-life" usage of medicines, where effectiveness of therapy is not just related to biological benefit but is also influenced by other issues such as the patient's adherence (e.g. problems with dosing regimen, adverse effects); PPA assesses only the effect of the intervention in those who adhere perfectly to the protocol (i.e. the "ideal" patient) and is said to reflect the innate efficacy of the intervention.[8]

Table 4: Commonly used statistical terms[8,10,37-39]

TERM	EXPLANATION
Null hypothesis	A test of significance, which is based on the premise that the treatments being compared are equally effective (i.e. the reverse of what the study is designed to show). The aim is to reject this null hypothesis, thereby showing statistical significance (suggesting that a true difference exists)
Power of a Test	A measure of how likely it is to be able to find a certain size of difference between the groups being compared, assuming such a difference exists, e.g. many studies use a power of 80% and a significance level of 5%
Sample size	This refers to the desired number of subjects in a study. It is calculated with reference to the power of the study, the significance level chosen for the study and the departure from the null hypothesis (the so-called minimal effect size/clinically important difference in outcome between the two treatments) and the study design
Intention to treat analysis (ITT):	An analysis where all of the participants who entered an interventional study are included in the results, whether or not they took or correctly received the intervention to which they were allocated, i.e. the way it was intended to treat subjects, not the way in which they were actually treated. Therefore dropouts are still included in order to maintain comparability between groups
Per Protocol Analysis (PPA)	Analysis of results of an interventional study based on the subset of subjects who complied with the protocol
p-value	The probability of how likely a particular result in a study occurred by chance alone if the null hypothesis were true. p<0.05 means less than 5% (1 in 20) probability, which is the conventional level of statistical significance. Also referred to as Type 1 error
Confidence Interval (CI)	A range of values around a study result within which, at a given level of confidence, the true population value is likely to be found (e.g. 95% CI means 95% confidence that the true population value lies within the range of the confidence interval presented)
Relative risk (RR)	The risk of an outcome in subjects with a particular characteristic (e.g. treatment) compared with the risk of that outcome in subjects who don't have that characteristic. It is calculated by dividing the rate of the event in one group of patients in study by the rate of events in the comparator group. A RR of 1 indicates no association between treatment and outcome; RR >1 indicates a positive association between treatment and outcome; RR <1 indicates a negative association between treatment and outcome
Absolute Risk (AR)	Chances of something happening in a specific population, i.e. number of events in a population in a time period, divided by the total population at the start of the time period
Absolute risk reduction (ARR)	The amount that by which a treatment reduces the risk of an event
Odds Ratio (OR)	The ratio of odds of outcome (e.g. disease) occurring in a group exposed to a possible risk factor compared to the odds for a non-exposed group. An OR of 1 indicates no association between exposure and outcome
Hazard ratio (HR)	The ratio of the chance of an event occurring in one study arm of a comparative study, compared to the chance of that event occurring in the other study arm
Number needed to treat (NNT) (or harm [NNH])	Number of people needed to be treated in order to achieve benefit [or develop an adverse outcome] in one person (it is calculated by dividing 1 by the ARR)

Statistical tests: the presentation and interpretation of results depend on the study design and the statistical

tests used to analyse the data. Although there are many statistical tests available, **many studies use a small number of statistical tests**; in most cases the p-value and confidence intervals are frequently used (see table 4). If unfamiliar statistics have been used in the analysis, the validity of such usage should be questioned. In addition, it is important to know if the statistical tests used in the study were part of the original statistical analysis plan (**a priori** analysis), or were adopted after the study protocol was finalised (**post-hoc** analysis); the latter analysis lessens the scientific validity of results.[38]

The **p-value** relates to the probability that any particular outcome would have arisen by chance.[38] Standard scientific practice usually deems a p-value of less than one in 20 (written as $p<0.05$) to be statistically significant;[37,38] the smaller the p-value, the lower the likelihood that the result happened by chance, and the more certain that there is a difference between the two treatments being compared.[37]

Confidence intervals (CI: usually 95% CI is reported) can be calculated for most statistical tests and provide information about the magnitude of the result, which is useful in evaluating the clinical relevance of the results.[38] The **larger the sample size the narrower the CI,** which increases the **precision** (i.e. accuracy) of the result.[38] Statistical significance can also be inferred from CI as follows: if the 95% CI for a clinical trial comparing mean outcomes between two different treatments does not include zero (which reflects the null value or no difference between the two treatments), statistical significance is implied and a p-value <0.05 is assumed; if 95% CI includes zero, this result is taken as non-significant i.e. lack of evidence to support a difference between the treatments (p-value $\geq$0.05).[38]

Risk and benefit can be assessed using **relative risk** (RR) when dealing with a full study "population" as in a cohort study or clinical trial, or **odds ratio** (OR) when the full "population" is unknown as in a case-control study (see table 4).[37,38] Statistical significance can be inferred from the CI if the 95% CI for RR or OR includes one (as one represents no difference in risk between groups).[38] RR estimates do not take into account the individual's baseline risk of achieving the intended outcome without the intervention, and tend to overestimate the benefits of an intervention.[40] **Absolute risk** (AR) estimates (see table 4) reflect the baseline risks and are better at discriminating between large and small treatment effects.[40]

Hazard Ratio (HR) is the ratio of the chance of an event occurring in one study arm of a comparative study, compared to the chance of that event occurring in the other study arm.[39] HR may be used to present results involving survival or time-to-event data (e.g. mortality, acute myocardial infarction) and is frequently reported alongside a measure of time.[39]

HOW TO INTERPRET THE RESULTS AND DISCUSSION SECTIONS

Usually, the results section is presented in a series of graphs / tables and explanatory text which should enable the reader to understand the results and their potential relevance to clinical practice. In a RCT, it is important to review the **baseline characteristics of subjects** in each arm as any difference might impact on the interpretation of the findings; similarly, it is important to check that all study subjects were followed up and accounted for in the study.[6] Another important aspect of any paper is the discussion of **potential limitations of the study;** limitations include problems with inclusion / exclusion criteria, insufficient numbers recruited and loss to follow-up of patients.[25] Any of these issues may impact on the (**internal**) **validity** of the study findings (i.e. **whether the results are true or may be affected by bias**).

HOW TO CHECK THE APPLICABILITY OF A STUDY

Protocols for research studies, especially clinical trials, normally contain strict inclusion / exclusion criteria to ensure a homogeneous study population for each arm (e.g. the only difference between the two arms of the trial is the intervention) and in this way increase the scientific validity of the results. However, **this may adversely affect the relevance of the study** (known as **external validity**) for specific patient groups.[41] Therefore each healthcare provider needs to check the applicability of published research for his/her clinical practice. Table 5 proposes a checklist for assessing the external validity of a clinical trial in published scientific papers.

Table 5: Checklist for assessing the external validity of a clinical trial[42]

- Setting of the trial e.g. healthcare system, country, recruitment from primary or secondary care
- Selection of patients e.g. eligibility criteria, exclusion criteria
- Characteristics of randomised patients e.g. baseline clinical characteristics, ethnicity, severity of disease and comorbidities
- Outcome measures and follow-up e.g. clinical relevance of outcomes, frequency of follow-up, adequacy of the length of follow-up
- Adverse effects of treatments e.g. rates of discontinuation of treatment, exclusion of patients at risk of complications or who experienced adverse effects during a run-in period, intensity of trial safety procedures

SUMMARY

When reviewing the medical literature, it is important that healthcare professionals check that (1) the study design is the most appropriate to answer the specific question being investigated, (2) the study is methodologically correct and (3) the results are relevant to the individual patients under their care. Table 6 provides a summary checklist of relevant questions in the review of a clinical paper.

Table 6: Summary checklist of questions to address when reading a clinical paper[6]

General	Validity	Results	Applicability
- Clear question (information on PICO* provided)? -Appropriate study design?	Correct: - Randomisation - Sample size - Study conduct	Correct: -Presentation -Statistics Results clinically relevant?	Study population and results relevant to your practice?

***PICO**=**P**opulation of interest; **I**ntervention; **C**omparator; **O**utcome

USEFUL SOURCES OF INFORMATION

www.casp-uk.net/. CASP (Critical Appraisal Skills Programme) which provides checklists and other tools to assist critical appraisal

www.bmj.com. Series of articles entitled "How to Read a Paper" – search under **Trisha Greenhalgh** as author

www.consort-statement.org/. This comprises a checklist and a flow diagram outlining the minimum recommendations for reporting RCTs

www.gov.ie/en/collection/c9fa9a-national-clinical-guidelines/ National guidelines published by the National Clinical Effectiveness Committee (NCEC)

List of references available on ePublication on www.nmic.ie.
Date of publication: March 2022

References for How to review a clinical paper (28th March 2022)

1. Sackett et al, Evidence based medicine: what it is and what it isn't. BMJ 1996; 312: 1-2
2. Djulbegovic B, Guyatt G, Progress in evidence-based medicines: a quarter century on, Lancet 2017;390:415-23
3. Murad M et al, How to read a systematic review and meta-analysis and apply the results to patient care. JAMA 2014; 312: 171-9
4. What is clinical effectiveness. Clinical Effectiveness Unit, Dept of Health, Ireland. Available online at: http://health.gov.ie/clinical-effectiveness-unit/. Accessed 13 January 2022
5. National Clinical Guidelines. Clinical Effectiveness Committee, Dept of Health. Available online at: https://www.gov.ie/en/collection/c9fa9a-national-clinical-guidelines/ Accessed 13 January 2022
6. NMIC - How to read a Clinical Paper NMIC 2016 Volume 22 Number 3 (available on file)
7. Getting your bearings: what is this paper about? in "How to read a paper: the basics of evidence-based medicine" Sixth Edition. Editor: Trisha Greenhalgh. Publishers: Blackwell publishing UK 2019. Chapter 3, pages 29-43
8. Anonymous, Understanding statistical terms: 3, Drugs & Therapeutics Bulletin 2009; 47(5):59-60
9. Grimes DA, Schulz KF, An overview of clinical research: the lay of the land, Lancet 2002;359:57-61
10. Anonymous, Understanding statistical terms: 1, Drugs & Therapeutics Bulletin 2009; 47(2):22-36
11. Assessing methodological quality, in "How to read a paper: the basics of evidence-based medicine" Sixth Edition. Editor: Trisha Greenhalgh Publishers: Blackwell publishing UK 2019. Chapter 4, pages 43-61
12. Anonymous, Understanding observational studies, Drug and Therapeutic Bulletin 2016;54(9):105-108
13. Grimes DA, Schulz KF, Cohort studies: marching towards outcomes, Lancet 2002;359:341-45
14. Growing up in Ireland. Available online at: http://www.growingup.ie/index.php?id=9 Accessed 11th February 2022
15. The Irish Longitudinal study on ageing. Available online at: www.tilda.tcd.ie Accessed 11th February 2022
16. Grimes DA, Schulz KF, Case-control studies: research in reverse, Lancet 2002;359:431-34
17. Jick H et al, Risk of idiopathic cardiovascular death and nonfatal venous thromboembolism in women using oral contraceptives with differing progestagen components. Lancet. 1995; 346(8990):1589-93.
18. Grimes DA, Schulz KF, Descriptive studies: what they can and cannot do, Lancet 2002;359:145-49

19. The EUROASPIRE survey of cardiovascular prevention and diabetes in 24 countries in Europe. European Heart Journal (2015) 36, 950–955
20. Healthy Ireland Survey 2021. https://www.gov.ie/en/publication/9ef45-the-healthy-ireland-survey-2021/ Accessed 11th February 2022
21. McCormick PA et al, COX-2 inhibitor and fulminant hepatic failure. (Letter.) Lancet 1999; 353: 40-44
22. McBride W. 1961. Thalidomide and congenital malformations. Lancet 1:358
23. Grimes DA, Schulz KF, Bias and causal associations in observational research, Lancet 2002;359:248-52
24. Papers that summarise other papers (systematic reviews and meta-analyses), in “How to read a paper: the basics of evidence-based medicine” Sixth Edition. Editor: Trisha Greenhalgh. Publishers: Blackwell publishing UK 2019. Chapter 9, pages 117-136
25. Anonymous, Understanding systematic reviews and meta-analyses. Drugs & Therapeutics Bulletin 2013; 51: 117-120
26. Anonymous, Understanding statistical terms: 5 – systematic reviews and meta-analyses, Drug & Therapeutics Bulletin 2009;47(8):95-96
27. Murad MH, Montori V, Ioannidis J et al, How to read a systematic review and meta-analysis and apply the results to patient care, JAMA 2014;312(2):171-179
28. About us: Cochrane. Available online at: http://www.cochrane.org/about-us accessed on the 3rd February 2022
29. Leen B, bell m, McQuillan P, Evidence Based Practice: a practice manual (Kilkenny 2014), Health Service Executive available online https://www.hse.ie/eng/about/who/qid/use-of-improvement-methods/nationalframeworkdevelopingpolicies/8-evidence-based-practice-a-practice-manual.pdf. Accessed 10th February 2022
30. McGlothlin A, Minimal Clinically Important Difference: defining what really matters to patients. JAMA 2014; 312: 1342-3
31. Stokes L, Sample size calculation for a hypothesis test. JAMA 2014; 312: 180-1
32. ICH E9 - Note for guidance on statistical principles for clinical trials. (CPMP/ICH/363/96), September 1998. Available online at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002928.pdf. Accessed 25th February 2022
33. Kaji A, Lewis R, Noninferiority trials – is a new treatment almost as effective as another? JAMA 2015;313(23):2371-2372
34. Connolly SJ et al, Dabigatran versus Warfarin in Patients with Atrial Fibrillation (RE-LY). NEJM 2009; 361:1139-51
35. Patel MR et al, Rivaroxaban versus Warfarin in Non-valvular Atrial Fibrillation (ROCKET AF). NEJM 2011; 365:883-91
36. Granger CB et al, Apixaban versus Warfarin in Patients with Atrial Fibrillation, (ARISTOTLE). NEJM 2011; 365:981-92

37. Anonymous, Understanding statistical terms: 2, Drugs & Therapeutics Bulletin 2009; 47(3):35-36
38. Statistics for the non-statistician, in “How to read a paper: the basics of evidence-based medicine” Sixth Edition. Editor: Trisha Greenhalgh. Publishers: Blackwell publishing UK 2019. Chapter 5, pages 62-78
39. Sedgewick P, Endgames - Interpreting hazard ratios, British Medical Journal 2015; 351:h4631
40. Akobeng AK, Understanding measures of treatment effect in clinical trials, Archives of Diseases in Childhood 2005;90:54–56
41. Moher D, Hopewell S, Schulz K, et al, CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trial, British Medical Journal 2010;340:c869
42. Rothwell P, External validity of randomised controlled trials: “To whom do the results of this trial apply?” Lancet 2005;365: 82–93