REVIEW PAPER

JMS Journal of Medical Science

Study designs in medical research and their key characteristics

Jan K. Nowak
Department of Paediatric Gastroenterology and Metabolic Diseases, Institute of Paediatrics, Poznan University of Medical Sciences, Poland
https://orcid.org/0000-0003-0953-2188

Jarosław Walkowiak
Department of Paediatric Gastroenterology and Metabolic Diseases, Institute of Paediatrics, Poznan University of Medical Sciences, Poland
https://orcid.org/0000-0001-5813-5707

Corresponding author: jarwalk@ump.edu.pl

DOI: https://doi.org/10.20883/medical.e928

Keywords: research methodology, study design, randomized controlled trial, cohort, case-control, review, cross-section

Received 2023-09-11
Accepted 2023-12-22
Published 2024-01-03

How to Cite: Nowak JK, Walkowiak J. Study designs in medical research and their key characteristics. JMS. 2024 Jan. 3 [cited 2024 Jan. 31];92(4):e928.. doi:10.20883/medical.e928

ABSTRACT

Medical research study designs are many and varied. At first glance, they may be difficult to distinguish. Knowledge of their specific strengths and limitations is helpful for investigators planning new projects and for readers of the medical literature. The aims of the review are threefold: (i) to present an overview of medical research types, (ii) to attract attention to multiple characteristics of medical study designs, and (iii) to provide a concise educational resource for young researchers in the health sciences. Analysing the characteristics of medical study designs leads to achieving the goals.

1. Introduction

Designing a medical research project involves choosing methods to address a research question. The selection of the proper study design is critical for success and determines the limits for drawing reliable conclusions. Various types of studies have individual properties and can answer different types of questions [1]. This text briefly summarises the most common medical study designs. It also hopes to underscore several less known research design characteristics within the traditional division of study types. We refer the reader to recent reviews for novel proposals on how medical research design can be transformed or viewed from the specific perspective of personalised medicine [2,3].

2. Major types of research studies in biomedicine

The diversity of major study designs [4,5] highlights differences in their strong and weak points. While a cross-sectional study may be used to investigate disease prevalence, a case-control study can identify its risk factors. Furthermore, the cohort study may trace the disease course, and an interventional trial can verify if a proposed treatment works.

Of note, all of these studies are at risk of selection bias, which means that the study results can be applied (generalised) to the population of which the study group is representative. Consequently, the results of a study conducted on neonates cannot be used to conclude the treatment of older patients. Likewise, conclusions from research done only in women may not apply to males.

2.1. Observational studies

2.1.1. Cross-sectional study

Research of this type collects information about several characteristics of individual study participants at one time point (which does not need to be the same for each person). Then, the characteristics are summarised and compared between groups or relationships to identify potential connections. Both methods can be used, too. For example, medical students' final exam results in biophysics can be analysed in the context of their physical activity. physical activity is the exposure, and test results are the outcome. The main advantage of the cross-sectional study is an analysis of many exposures (e.g., risk factors) and outcomes (diseases, parameters). They are also cost-effective.

The major limitation is confounding, which may occur in any study. Yet, the cross-sectional design accounts for it. Confounding occurs when multiple factors coexist, and we do not know which is important. Thus, we may assume that healthy people are more physically active and often eat healthy.. However, more information is needed to say if physical activity results from health, health from activity, or both, depending on the diet. Consequently, a cross-sectional study rarely establishes cause-and-effect relationships. It is much better used to screen for correlations or to rule out strong effects.

2.1.2. Case-control study

The study focuses on a single outcome in this design. For example, the outcome can be the presence or absence of an inflammatory bowel disease. Patients with the disease and healthy controls can be enrolled to compare their past (usually), present, and future characteristics or either of the features.. Control participants can be matched to cases for age and other factors to minimise confounding. If the individuals are of the same age, then it is unlikely to affect the differences between the groups (though this issue has additional layers of complexity [6]). More controls than patients are often recruited to boost the study's statistical power. The result of the case-control study typically is an odds ratio, which is different from the more intuitive risk ratio (see next subsection). A more advanced view of the case-control study requires a recognition of the dynamic nature of the population [7]. The case-control design requires a high degree of organisation, and large-scale case-control studies frequently rely on long-established information technology systems. The case-control design also suffers from recall bias if information about the participants' past is obtained; people with and without disease may remember their past differently. Suppose adequate medical records and systems are available. In that case, this study enables a low-cost investigation of the relationship between a wide range of exposures and the study outcome, such as disease.

2.1.3. Cohort study

Time is the central concept in a cohort study. In a cohort study, many individuals are monitored over an extended period to observe the development of specific health effects, such as asthma, in a large group of children followed up for 20 years. Data about exposures and outcomes are often collected at different intervals throughout the study. A project may extend to the long-term storage of biological samples. The cohort study establishes a link between a specific exposure (such as detergent use) and later outcome (such as asthma), usually expressed as a relative risk (or risk ratio). However, such findings also need more proof of cause and effect: the cohort study is subject to confounding, even if it may be easier to mitigate than in other study types. As one cohort study may analyse multiple exposures and outcomes, extensive cohorts have been established and thus provided a wealth of data on risk factors for lifestyle diseases. Cohort studies are often large (thousands of participants). These studies offer the highest quality of evidence when the prospective design is used: starting observation of the group after all the questionnaires are ready ensures that the data are complete. Information from existing medical records or databases, i.e.,

a retrospective study, bears the risk of missing data. However, the retrospective cohort design is sometimes applied when patients with a rare disease (such as cystic fibrosis) are followed up for many years in the same centre. **Figure 1** summarises the typical timing for major types of observational studies.

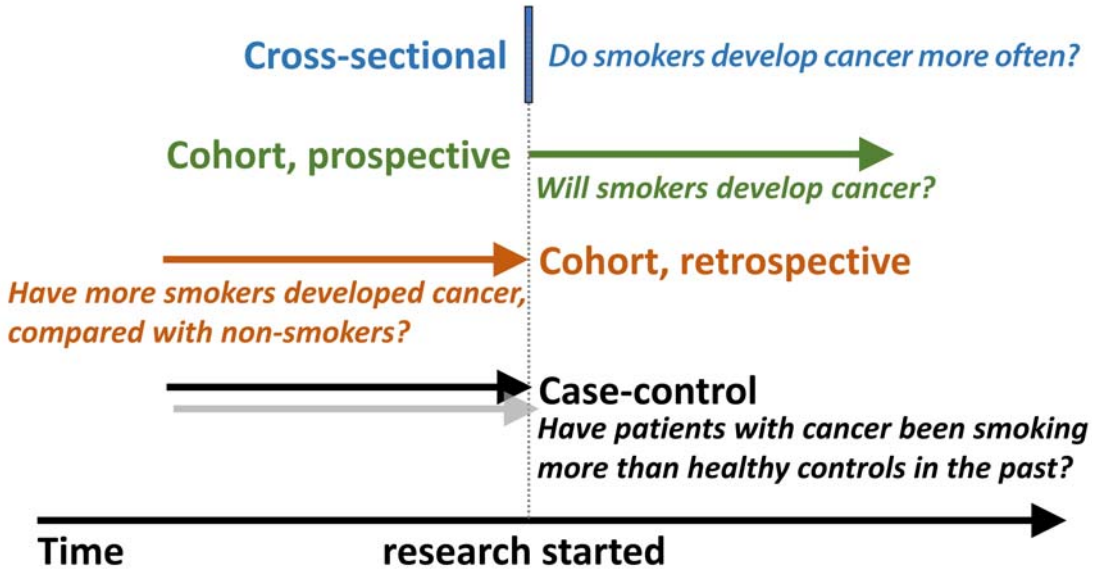

Figure 1. The relationship between major types of observational studies and time determines the types of questions that they can address. The vertical dotted line indicates the time the research was carried out. Many modifications are possible, including prospective case-control studies.

2.1.4. Studies of diagnostic accuracy

Assessment of a medical test's diagnostic value is a frequent research subject, often categorised separately from medical management (observational). Evaluating diagnostic accuracy requires an established test of reference, the golden standard. Comparing and assessing new methods can be done by comparing them with the reference method. Studies of diagnostic value are also subject to bias resulting from the selection of the investigated group of patients because the diagnostic value does not need to be identical in people with two different diseases. An example of a diagnostic value study is assessing of a new, non-invasive test for diagnosing early-stage lung cancer. The latest trend in diagnostic value research is embedding such studies within large interventional trials, enabling early biomarker discovery. Results from this type of research are often presented using specific metrics (such as sensitivity, specificity, positive/negative likelihood ratio, positive/negative predictive value, and the AUC – area under the receiver operating characteristic curve) [8]. Moreover, statistical inference can compare such results between various tests to establish significant differences.

2.1.4.1. Specific issues related to studies of diagnostic value

All the above-mentioned metrics of diagnostic value are useful but require careful interpretation. A key question is whether the studied group is representative of the population intended for the study results application. Sensitivity and specificity depend on the selected cut-off (threshold), which is essential yet seldom attracts attention. Sometimes sensitivity is more desirable (not to miss sepsis), while in other scenarios, specificity is required (evading false positive results in populational screening). Adapting the test for a scenario of use can often be done with the right threshold.

Studies of diagnostic value frequently report the AUC, which reflects whether the test has high sensitivity and specificity for the same threshold. Notably, an AUC of 0.5 means zero diagnostic value and an AUC of 1 suggests perfect dis-

crimination between the groups. AUC 0.7 to 0.8 is considered acceptable, 0.8 to 0.9 is considered good, and AUC above 0.9 is considered excellent. Even if the AUC is high, the test may not be good, especially if almost all analyzed data are negative or positive (group imbalance), because in such a setting, even tossing a coin would give a high AUC. AUC has also been used to summarise the performance of prognostic regression models or, more recently, artificial intelligence-derived diagnostic classifiers, but they are subject to the same limitations. Awareness of these pitfalls helps to interpret the results of diagnostic value research. Moreover, searching for other studies that show the same effects (independent replication) is useful.

2.2. Interventional studies

2.2.1. Randomised trial

Randomization is core to the design of interventional trials because it protects against confounding. Therefore, it enables the researcher to demonstrate cause-and-effect relationships. Random allocation of participants to a new intervention (usually medication) or neutral substance (placebo) maximizes the chances that receiving the intervention or placebo is the only difference between the two groups. Then, any observed differences can be ascribed to the intervention.

An important part of randomized study design can be using a neutral substance (placebo) or fake intervention. It removes the effect of a positive attitude brought about by awareness of a medical intervention (the placebo effect). The study is blinded if the patient does not know what intervention they received. Suppose the researchers do not know which intervention patients receive either. In that case, it is double-blinded. The method protects against assigning higher patient scores on the treatment (that scientists could prefer to show that the intervention works).

Much attention is given to the number of patients enrolled in a randomized clinical trial, how they are assigned to each group, and how many complete the study. If too many patients withdraw from the study, it may indicate a problem not predicted from the start, such as severe adverse reactions. High exclusion rates at the start of the study suggest stringent inclusion criteria or organizational problems.

There are two main ways to approach the interpretation of results from a randomized clinical trial: intention-to-treat and per protocol. In intention-to-treat, the groups are compared based on the initial assignment. If someone was prescribed the investigated intervention but had side effects and withdrew, this person will still be included in the intention-to-treat analysis. The results of such analyses are useful for physicians and insurance agencies because they link treatment effects to the prescription of intervention. Per protocol (on-treatment) analysis compares groups of patients who received study and control interventions until the end of the study. Therefore, the groups are smaller, and the comparison ignores that many patients were excluded. This analysis is interesting as it may be more sensitive to some effects of interventions but has an increased risk of bias.

Notably, the interventional study, especially the randomized controlled trial, bears higher ethical and legal requirements than the observational study, which usually incurs high financial costs. Randomized trials are registered in appropriate databases (such as clinicaltrials.gov) before they start and thus: (i) the study cannot be manipulated, (ii) it is known the study was attempted even if results are not published, (iii) other teams do not start the same expensive study that someone else is already doing. Teference textbooks summerise the clinical trial methodology [9,10]. Organizational and regulatory aspects of clinical trials are of great importance.

2.2.1.1. Randomization

The process of randomization frequently involves the use of computer tools. Block randomization typically uses pre-generated blocks of a specific size to assign patients to an intervention. If the block size is two, there are two possible blocks (orders): placebo-intervention and intervention-placebo. The randomization list consists of subsequent blocks that determine the order of group allocation [11]. This approach has a limitation: if the block size is small and the researcher knows that one patient received a placebo, she/he will be aware that the next person will be given the investigated drug. Larger block sizes disguise the intervention allocation more efficiently. However, then the group sizes may be uneven in the end. In randomized studies, study groups might

not be perfectly equal, as this results from randomization itself (unless designed otherwise on purpose). A possibility to maximize the similarity of the compared groups in confounding factors (such as sex) by stratification also exists and can be done by constructing two separate randomization lists: if the patient is female, the following allocation (placebo or intervention) is done according to the list for women (and not from the list for men). With such a setup, the study and control groups will not differ in the proportion of women. Stratification can be done with more factors, such as age, recruitment centre and medication use. Many randomized studies are multicenter, as it is often too difficult to recruit enough patients in one city and because some patients are lost to follow-up. Of note, "Mendelian randomization" does not relate to clinical trials but is an observational method to establish cause-and-effect relationships from genetic and epidemiological data analysis.

2.2.1.2. Outcomes and statistics

The time and intervention aspect of a randomized trial makes it possible to analyze trial data in multiple ways. One comparison method involves defining outcomes that can be compared at the end, such as the size of a tumour between two groups. Another approach is to check which group experiences slower tumour growth (analyzing a delta value equal to tumour size after minus tumour size before). There are other ways to approach clinical trial data, and there are also several statistical approaches and specific outcome types, such as survival, which is a complex issue. From a statistical standpoint, it is essential to appreciate one fact related to clinical trial design and sample size calculations.

In most cases, it will be challenging to statistically prove an effect of intervention when the principal outcome is defined as a binary variable (1 or 0, Boolean), like whether someone achieved a response or not. Researchers use a statistical measure called the p-value to determine if an intervention is effective. Typically, the standard assumptions for the p-value threshold (alpha or type I error) are set at 0.05 and the power at 0.8 (also known as beta or type II error at 20%). Suppose we want a statistically significant result in a trial where the first group has a 35% success rate and the second group has a 40% success rate (a 5% absolute difference and a 14% relative difference). In that case, we must recruit almost 3000 participants.. A more significant 10% absolute difference between the groups (first group 30% vs. second group 40%) would reduce the required sample size to 700 participants, which is still a large number. It is usually easier to prove the effect at a moderate sample size by measuring a continuous variable (such as blood pressure).

Sample size calculations for clinical trials are discussed by DELTA[2] guidelines (Difference ELicitation in TriAls) [12]. Surrogate outcomes are used because the most pertinent health-related outcomes are binary and because they take very long to develop (like death or stroke). Surrogate outcomes are known to be health-related, but they are easier to measure than "hard" endpoints such as mortality. Of course, endpoint use translates to what the study means for the patient and the physician, who may explain, depending on what the trial measured: "This pill will reduce your risk of death in the next ten years by 5%" or "This pill will normalize your blood pressure, which will probably help you live longer." Apart from surrogate outcomes, composite outcomes are used, where the occurrence of any endpoints is counted identically, be it death, stroke, or hospitalization. All will agree that death is not the same as hospitalization, and the two will not occur with the same frequency in the trial. Therefore, the adequate use and interpretation of composite outcomes is a challenge [13]. They sometimes need to be used because the most critical investigated outcomes occur rarely, and it is impossible to carry out extensive and long enough trials to obtain the answer. Apart from the primary outcome, randomized trials often have multiple secondary outcomes. The selection of appropriate outcomes in various diseases is not straightforward and constitutes a strand of research that quickly gains interest [14].

2.2.1.3. Commercial aspects

Many clinical trials are non-commercial, focusing on already registered medications or non-pharmaceutical interventions, such as surgery, diet, or exercise. However, randomized trials are also crucial for regulatory approval of new pharmaceuticals or medical devices. Therefore, the industry funds many and attracts additional

scrutiny because of the involved financial interests. Essential regulatory requirements accompany such trials but do not eliminate the conflict of interest risks [15].

These risks remain difficult to appreciate for the reader, as evidenced by a recent analysis of undisclosed payments to medical scientists [16]. Moreover, researchers only sometimes understand the conflict of interest similarly [17] because the related problems are diverse and not equally perceived in all cultures. Activities of the tobacco industry provided numerous and varied examples of how researchers can be influenced to evade presenting complete information, draw attention to insignificant topics or manipulate the reader and the public opinion [18]. A scientist receiving payments from a company (or owning its stock) may present biased views. Therefore, it is wise to read the medical literature critically, not ignore "conflict of interest" sections, consider authors' affiliations, and also keep in mind that study results or conclusions are often reported in misleading ways (with a "spin") [19]. However, the last of these issues overlap with a general trend towards more usage of positive language in the research literature [20]. In brief, a critical approach is always valuable while reading any research literature, including this text.

2.2.1.4. Phases of clinical trials

Table 1 lists the four main phases (I-IV) of clinical trials. Phase 0 studies may also be carried out to assess drug bioavailability and metabolism in healthy participants. Group size depends on the phase, ranging from below 100 participants (Phase I) and 100–300 patients (Phase II) to between 300 and a few thousand (Phase III).

Early Phase I investigates how the body reacts to a new substance. Phase I study carried out in healthy volunteers, defines optimal dosage and checks for side effects. Phase II investigation is a larger trial in patients that searches for evidence of efficacy and extends the safety assessment. Phase III trials are fundamental investigations of efficacy and safety and are the primary source of safety information available to physicians. Therefore, the significant difference between Phase II and Phase III is that it is still being determined if the intervention works before Phase II. Phase III studies are also more extensive and provide more precise results than Phase II. Phase IV studies are done after medication is approved and may include evaluation of treatment effects and post-marketing safety surveillance.

2.2.1.5. Simulation and in silico studies

In-silico studies help pre-clinical research, that is, computer modelling (computer processors are made of silicon).. Tools are available to predict ligand binding dynamics and off-target effects, facilitating the development of safer pharmaceutical compounds, which are then tested biologically.

Simulations may also be helpful in later phases of clinical research. Big data from electronic health records can be used to conduct simulated randomized controlled trials [21]. Such studies do not replace clinical research, but they may be used to provide additional information about the relationships between understudied population characteristics and the efficacy or safety of interventions. The primary example is ethnicity, especially when a randomized controlled trial does not reflect the entire target population well. Simula-

Table 1. Main phases of clinical trials. The duration of studies is between several months for Phase I and a few years for Phase III studies, which are usually done in multiple centres to recruit a sufficient patients.

Phase of clinical trial	Focus
Phase I	– Choosing optimal dose – Showing safety – Healthy participants or patients with cancer
Phase II	– Showing that intervention works – Confirming safety – Done in patients
Phase III	– Confirming and measuring the efficacy precisely – Confirming and assessing safety in detail – Done in patients
Phase IV	– Monitoring safety in real life after the intervention becomes broadly available – Done in patients usually prescribed the intervention

tion can attract attention to phenomena that are otherwise overlooked and may be important for specific subgroups of patients. Such computer methods (trial emulation) are complex [22].

2.2.2. Non-randomized interventional studies

Predominant interventional studies in biomedicine are now mostly randomized, placebo-controlled trials. However, some other interventional designs warrant discussion. One is the before-after design, applicable in uniformly progressive illnesses and when treatment is expected to yield outstanding benefits, such as gene therapy for rare diseases. Non-randomized investigation of various interventions in a real-life setting may be helpful to establish efficacy and broaden the scope of safety assessment. Some of the non-randomized interventional studies do not include a control group. Unfortunately, these quasi-experimental designs are prone to bias, so their results should be interpreted cautiously [23].

2.3. Review studies

2.3.1. Narrative review

A narrative review provides knowledge summarized by experts in the field without using a systematic methodology. Reviews of this type authored by most experienced researchers are often most helpful. Such reviews may serve as works of reference, similar to textbook chapters.

2.3.2. Systematic review

A systematic review aspires to synthesize all the information on a given topic. The goal is achieved through the application of systematic search methodology. A query is constructed and used to search literature databases. The identified articles are assessed for relevance to research questions. Quantitative information from these publications is summarized. A report from a systematic review presents the search strategy, the keywords used, the fields searched, and the databases used. Specific inclusion and exclusion criteria increase the quality of the systematic review (and the meta-analysis – see below).

2.3.3. Systematic review with meta-analysis

Introducing meta-analysis to a systematic review requires specific qualitative summary methods [24,25]. Studies included in meta-analysis are assessed for various forms of bias. The data available from publications and research teams are combined to obtain one more reliable summary metric of effects.

Reviewing study quality (dealing with bias) is integral to the meta-analysis process. The risk of bias is determined using appropriate tools. For randomized controlled trials, this can be Cochrane's risk-of-bias assessment tool 2 (RoB 2, [26]), which refers to potential limitations in randomization, studied intervention, dealing with missing data, approach to outcome measurement and result presentation. Other tools which can be used to assess the quality of observational studies include checklists from the National Institutes of Health or the Scottish Intercollegiate Guidelines Network (SIGN). SIGN guidelines include the quality assessment tool for diagnostic accuracy studies (QUADAS-2), which relates to patient selection bias, blinded interpretation of tests, adequate gold standard (reference) measurement, and exclusions of cases from analysis [27]. Although designed to help with meta-analysis, these tools are also valuable when planning or starting research because they guide how to conduct studies well.

There are different statistical approaches to meta-analysis, depending on the main research question. It is possible to synthesise many studies with univariate meta-analysis and compensate for study characteristics using meta-regression, which may be especially useful when data from individual patients are available to the researcher. There are also other methods which allow for linking multiple characteristics to multiple effects and for indirect comparisons (network meta-analysis).

Clinicians are often interested in the results of a systematic review with meta-analysis as this type of article provides the highest quality of evidence. The methodology of meta-analysis itself, therefore, needs to be strict; moreover, registration of a systematic review with meta-analysis is often encouraged to prevent other teams from spending effort on the same topics. More primary data is needed in many crucial areas to draw significant conclusions in meta-analyses.

2.3.4. Guidelines and consensus reports

We list guidelines under reviews, even though they are a separate article type. The reason for

this is the vital role of systematic search in determining the optimal medical management. After identifying and reviewing adequate references, the experts propose a set of recommendations to guide clinical practice. Most commonly, the quality of evidence to address a specific question is labelled, and voting determines what the panel will suggest as best care. However, the strength of recommendations does not rely on the quality of evidence alone, and the direct practical knowledge of the current medical practice often influences recommendations. Many guidelines cite hundreds of works; commonly, the number of referenced articles is a few times larger than in a regular review article.

2.4. Qualitative studies

Qualitative methods involve, among others, interviews, focus groups, experiments, observation, analysis of documents, and secondary sources. Qualitative and quantitative methods can be mixed [28]: e.g., interviews may be used to identify main problems, and then a survey may characterize the issue statistically. The value of qualitative studies lies in their capacity to uncover perceptions, attitudes, and mechanisms that would remain difficult to capture by quantitative research alone. They focus more on the comprehensive experience, allowing detailed and rich insights as flexible questions permit. Qualitative research makes it easier for the researcher to go beyond her/his preconceived ideas but requires much work, even to characterize small groups.

2.5. Case reports

Case reports present valuable information for learning, which is achieved through presenting typical or atypical courses of disease or unusual findings that may fuel further discovery. Case reports help identify therapeutic interventions' adverse effects and describe new diseases or risks. However, much of the information included in case reports may not be representative because of biological variability and ethnic, economic, and cultural differences, which give rise to bias. Case reports' conclusions often cannot be easily generalized, and the reports can be easily over-interpreted. The difference between a report of a series of patients and a cross-sectional or cohort study may take time to establish.

2.6. Basic biomedical research studies

This fundamental research, often involving models and experimentation [29], is behind most breakthrough discoveries fueling medical progress. Defining basic biomedical research is not straightforward, but the most commonly cited characteristic focuses on growing knowledge of nature (disease) and its mechanisms (without a specific application in mind). A cross-sectional study can meet basic research criteria when it focuses on understanding disease pathophysiology (and not necessarily treatment). It is easier to see that studies mainly involving animal models and advanced biomedical or molecular biology techniques represent basic research. Medics may overlook these pure research studies and their value because of inaccessibility (complexity) and the need for an obvious connection to medical practice. Indeed, such work is almost always of no direct value to the physician, even if it holds much promise for helping patients.

3. Discussion: characteristics of medical research

This article briefly introduced the main types of scholarly output in medical research. The reasons why specific research designs are more common than others may relate to the ease of conducting research under certain circumstances and the need to address specific questions. Considering the theory itself, we may see several dimensions that characterize these studies and the understanding of which may help produce unconventional study designs. Below, we list some aspects of biomedical research that are worthwhile considering while planning or appraising medical research.

The distinction between basic, translational, clinical, epidemiological, and applied research seems more related to the choice of methods than the study design itself. Like any tool, a study design can be used for different aims. Thus, both observational and interventional studies can be employed to achieve the goals of basic or clinical research. Most commonly, qualitative methods are utilized, but quantitative and mixed methods are also sometimes applied. Studies can be done only with original data, data from other projects, and both types of data sources (primary

and secondary), as is more and more commonly seen because of the availability of big data. The aspect of time further divides studies into prospective and retrospective, depending on how samples or data were gathered. Another distinction is between studies investigating, finding, and reporting new discoveries (new ideas, phenomena), proposing original solutions or hypotheses, and validation research. A prominent characteristic of the primary vs. validation research divide is between studies with small and large group sizes. The study size, in turn, is often associated with carrying out work in more than one centre or even across countries or continents. Finally, the research questions can concern various areas: determining disease frequency, understanding its pathophysiology, biomarker or treatment discovery, patient perspectives, and economic aspects. The choice of research questions and study design is pragmatic, reflecting maximum possibilities at a given budget, given the state of knowledge. Patients are also more and more commonly involved in designing research.

Medical research study design is often not a purely linear process, wherein the research question alone would determine the aims, study design, and methods. Various factors affecting individuals, groups, and consortia constrain research. Recognizing this interconnection between the research environment and the research project translates naturally to optimizing choices about details of study design and methods. Thus, planning medical research depends on researchers' expertise, interests, employment and funding, institutional, organizational, and technological capacities, responsibility for other projects, educational or commercial activity, perception of challenges in the research field, patient values and cooperation, safety, ethical and legal issues, and life situation. When planning a study, there are many factors to consider regarding its feasibility and potential impact. The most apparent trade-offs involve balancing cost versus sample size, duration versus clinical relevance and statistical power. Additionally, there may be trade-offs between a study's administrative and financial capacities and its ability to prove a cause-and-effect relationship in a randomized trial. Overall, the choice of study design is not dictated solely by the research question but remains under the strong influence of mostly unmodifiable external factors.

Additional issues include the use of proper statistics, the adequacy of the control group, and a valuable and honest section on study limitations. All research studies have limitations, but these can be overcome with proper reporting guidelines to help peers evaluate findings accurately (**Table 2**).

While commencing any medical research and while reading reports from clinical studies, it is also crucial that the pre-test probability for main hypotheses is at least tentatively assessed. If the hypothesis is implausible, it is indispensable to remain sceptical even with statistically significant results.

Table 2. Reporting guidelines for major study designs according to Enhancing the Quality and Transparency Of Health Research (EQUATOR) Network (all the guidelines are available aat www.equator-network.org). Over 600 guidelines adapted for various specific types of research are available, including economic evaluation, Mendelian randomization, pre-clinical studies, and study protocols.

Study design	Reporting guidelines	Full title of the guideline
Observational: cross-sectional, case-control, cohort	STROBE	The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies
Diagnostic value	STARD	STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies
Randomized trial	CONSORT	CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials
Systematic review and meta-analysis	PRISMA	The PRISMA 2020 statement: An updated guideline for reporting systematic reviews
Qualitative study	SRQR	Standards for reporting qualitative research: a synthesis of recommendations
Case report	CARE	The CARE Guidelines: Consensus-based Clinical Case Reporting Guideline Development

The diversity of medical research designs reflects the complexity of investigated phenomena and the diversity of settings where researchers attempt to answer research questions. A reader who understands the main characteristics of clinical study designs will more easily identify the most trustworthy information in the medical literature.

Acknowledgements

Contributions
JKN: conceptualization, investigation, writing – original draft. JW: conceptualization, investigation, supervision, writing – review and editing.

Conflict of interest
JKN reports a grant from Biocodex Microbiota Foundation and consulting for Procter & Gamble outside of the submitted work. JW reports personal fees and nonfinancial support from Biocodex, BGP Products, Chiesi, Hipp, Humana, Mead Johnson Nutrition, Merck Sharp and Dohme, Nestle, Norsa Pharma, Nutricia, Roche, Sequoia Pharmaceuticals, Vitis Pharma, outside the submitted work, as well as grants, personal fees, and nonfinancial support from Nutricia Research Foundation Poland, also outside the submitted work.

Funding sources
JKN received financial support through a grant from the Polish National Science Centre (2020/39/D/NZ5/02720).

References

1. Song JW, Chung KC. Observational Studies: Cohort and Case-Control Studies. Plast Reconstr Surg, 2010;126:2234–42. doi: 10.1097/PRS.0b013e3181f44abc.
2. Subbiah V. The next generation of evidence-based medicine. Nat Med, 2023;29:49–58. doi: 10.1038/s41591-022-02160-z.
3. Superchi C, Brion Bouvier F, Gerardi C, Carmona M, San Miguel L, Sánchez-Gómez LM, et al. Study designs for clinical trials applied to personalised medicine: a scoping review. BMJ Open, 2022;12:e052926. doi: 10.1136/bmjopen-2021-052926.
4. Gamble J-M. An Introduction to the Fundamentals of Cohort and Case–Control Studies. Can J Hosp Pharm, 2014;67:366–72.
5. Mann CJ. Observational research methods. Research design II: cohort, cross sectional, and case-control studies. Emerg Med J, 2003;20:54–60. doi: 10.1136/emj.20.1.54.
6. Pearce N. Analysis of matched case-control studies. BMJ, 2016;352:i969. doi: 10.1136/bmj.i969.
7. Vandenbroucke JP, Pearce N. Case–control studies: basic concepts. Int J Epidemiol, 2012;41:1480–9. doi: 10.1093/ije/dys147.
8. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open, 2016;6:e012799. doi: 10.1136/bmjopen-2016-012799.
9. Piantadosi S. Clinical trials: a methodologic perspective. Third edition. Hoboken, NJ: Wiley; 2017.
10. Friedman LM, editor. Fundamentals of clinical trials. 5. ed. Cham Heidelberg: Springer; 2015.
11. Kang M, Ragan BG, Park J-H. Issues in outcomes research: an overview of randomization techniques for clinical trials. J Athl Train, 2008;43:215–21. doi: 10.4085/1062-6050-43.2.215.
12. Cook JA, Julious SA, Sones W, Hampson LV, Hewitt C, Berlin JA, et al. DELTA 2 guidance on choosing the target difference and undertaking and reporting the sample size calculation for a randomised controlled trial. BMJ, 2018:k3750. doi: 10.1136/bmj.k3750.
13. Heneghan C, Goldacre B, Mahtani KR. Why clinical trial outcomes fail to translate into benefits for patients. Trials, 2017;18:122. doi: 10.1186/s13063-017-1870-2.
14. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. Trials, 2012;13:132. doi: 10.1186/1745-6215-13-132.
15. Hakoum MB, Noureldine H, Habib JR, Abou-Jaoude EA, Raslan R, Jouni N, et al. Authors of clinical trials seldom reported details when declaring their individual and institutional financial conflicts of interest: a cross-sectional survey. J Clin Epidemiol, 2020;127:49–58. doi: 10.1016/j.jclinepi.2020.05.026.
16. Baraldi JH, Picozzo SA, Arnold JC, Volarich K, Gionfriddo MR, Piper BJ. A cross-sectional examination of conflict-of-interest disclosures of physician-authors publishing in high-impact US medical journals. BMJ Open, 2022;12:e057598. doi: 10.1136/bmjopen-2021-057598.
17. Østengaard L, Lundh A, Tjørnhøj-Thomsen T, Abdi S, Gelle MHA, Stewart LA, et al. Influence and management of conflicts of interest in randomised clinical trials: qualitative interview study. BMJ, 2020;371:m3764. doi: 10.1136/bmj.m3764.
18. Grüning T, Gilmore AB, McKee M. Tobacco industry influence on science and scientists in Germany. Am J Public Health, 2006;96:20–32. doi: 10.2105/AJPH.2004.061507.
19. Ito C, Hashimoto A, Uemura K, Oba K. Misleading Reporting (Spin) in Noninferiority Randomized Clinical Trials in Oncology With Statistically Not Significant Results: A Systematic Review. JAMA Netw Open, 2021;4:e2135765. doi: 10.1001/jamanetworkopen.2021.35765.
20. Yuan Z, Yao M. Is academic writing becoming more positive? A large-scale diachronic case study of Science research articles across 25 years. Scientometrics, 2022;127:6191–207. doi: 10.1007/s11192-022-04515-2.
21. Wedlund L, Kvedar J. Simulated trials: in silico approach adds depth and nuance to the RCT gold-standard. NPJ Digit Med, 2021;4:121. doi: 10.1038/s41746-021-00492-7.
22. Wang SV, Schneeweiss S, RCT-DUPLICATE Initiative, Franklin JM, Desai RJ, Feldman W, et al. Emulation of Randomized Clinical Trials With Nonran-

domized Database Analyses: Results of 32 Clinical Trials. JAMA, 2023;329:1376–85. doi: 10.1001/jama.2023.4221.
23. Cuello-Garcia CA, Santesso N, Morgan RL, Verbeek J, Thayer K, Ansari MT, et al. GRADE guidance 24 optimizing the integration of randomized and non-randomized studies of interventions in evidence syntheses and health guidelines. J Clin Epidemiol, 2022;142:200–8. doi: 10.1016/j.jclinepi.2021.11.026.
24. Ahn E, Kang H. Introduction to systematic review and meta-analysis. Korean J Anesthesiol, 2018;71:103–12. doi: 10.4097/kjae.2018.71.2.103.
25. Pigott TD, Polanin JR. Methodological Guidance Paper: High-Quality Meta-Analysis in a Systematic Review. Rev Educ Res, 2020;90:24–46. doi: 10.3102/0034654319877153.
26. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ, 2019:l4898. doi: 10.1136/bmj.l4898.
27. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med, 2011;155:529–36. doi: 10.7326/0003-4819-155-8-201110180-00009.
28. Östlund U, Kidd L, Wengström Y, Rowa-Dewar N. Combining qualitative and quantitative research within mixed method research designs: A methodological review. Int J Nurs Stud, 2011;48:369–83. doi: 10.1016/j.ijnurstu.2010.10.005.
29. Bespalov A, Michel MC, Steckler T, editors. Good research practice in non-clinical pharmacology and biomedicine. Cham, Switzerland: Springer; 2020.