Why a Research Study needs a Design?

B. I. Avan (The Human Development Programme, The Aga Khan University. Karachi.)
F. White (Department of Community Health Sciences, The Aga Khan University. Karachi.)

Abstract

Medical research requires as much attention to management as it does to science. Effective research depends first upon the thoughtful statements of purpose and objectives, and related research questions. These important initial steps in turn drive key decisions regarding study type, plans for analyses and competent implementation within bounded timeframes and budgets. The research endeavour materializes by selecting appropriate study types based on coherent research questions, appropriate data collection, and ultimately concludes by drawing inferences on the basis of properly planned analyses. This artide reviews these principles and inter-relationships
(JPMA 51:197;2001).

Introduction

Scientific research may be viewed as systematic efforts to improve knowledge and practices. Effective research is mostly a matter of good planning and execution and hardly ever a matter of luck, although Pasteur once said 'chance favours the mind that is prepared". The term "serendipity" was inspired by the Persian tale of the three princes of Serendip, who often made discoveries by chance[1] and there are indeed accidental findings in medical research. For example, historically many drug therapies have been discovered through trial and error, although over the past two decades we have entered a new era of designing drugs based on specific knowledge at the molecular level. Such history aside however, the more important fact is that most serendipitous findings in medical research occur in formal scientific settings managed by well-trained and experienced scientists. Serendipitous findings are clearly secondary to the products of well-designed projects that are based on intelligent questions and well defined objectives. Systematic research requires appropriate research type and competent implementation within a bounded tirneframes and budgets. Systematic research is never fruitless. Even a negative finding contributes to the advancement of knowledge.

Intelligent research therefore requires considerable attention to planning and design, perhaps reminiscent of the carpenter's adage: "measure twice, cut once!" A research protocol is a doctrine of detailed instructions based on scientific principles and rules of evidence, which guide the researcher to collect and analyze the data in a manner that eventually results in answers to specified research questions. Research design is also a process of transition, which transforms knowledge from theory to practicality. Research aims are materialized by selecting an appropriate study type and instruments, implementation through data collection and, ultimately by drawing inferences on the basis of properly planned analysis. While one could conduct research without organized effort and a blueprint, but this would be analogous to the driver who would attempt to reach a destination in the least possible time vithout a map. Even more fundamental is linking the design to the underlying questions: in management terms, if you don't know where you are going, any road will take you there! Research questions evolve on the basis of existing knowledge. Careful selection and implementation of the appropriate research design will achieve relevant answers. Research methodology therefore links current status to the next level of knowledge.

Moral: Luck is a residue of design Branch Rickey

Study Type

Research in medical science broadly falls under either descriptive or experimental types. The study

purpose leads to what could be the most appropriate study type for it. If the study purpose asks "what" questions with a problem to identify and describe the concept, then descriptive study types are appropriate. However, when study purpose asks "why" and "how" questions with a problem to establish causal relationships, then experimental study types are relatively more appropriate. Study types and their appropriate utility would be further discussed in forthcoming articles. All that needs to be mentioned here is that the specific label of the study type not only mediates the uniform set of the study principles to the scientific community, but also enables the researcher to apply appropriate sampling techniques, measurement methods and analysis.

In practice, how study types are assessed in terms of scientific rigor varies in relation to the question being asked. Furthermore, studies designed to answer one question may also help to answer other questions, although aware that the study may not have been designed with this purpose in mind, and may (under these circumstances) be less rigorous than otherwise. The Concise Oxford Dictionary defines rigour in several ways, including:

"logical exactitude" and "strict enforcement of the rules". In the context of scientific research, the assessment of rigour requires reference mostly to internal validity (Index and comparison groups are selected and compared in such a way that observed differences between them on the dependent variables under study may, apart from sampling error, be attributed only to the hypothesized effect being studied). By contrast, "relevance" is defined as "bearing on or having reference to the matter". In science, this is termed "external validity" (generalizability). A study is externally valid, or generalizable, if it can produce unbiased inferences regarding a target population. The evaluation of external validity usually involves much more subject matter judgement than internal validity. Both internal and external validity contribute to the overall validity of a study, or "study validity".

With these considerations in mind, the following list of common study types is presented in descending order of rigour in relation to their contribution to the assessment of efficacy and effectiveness of interventions. This order is not necessarily appropriate for assessing study types for the assessment of prevalence, clinical course and prognosis of disease, clinical errors and the accuracy of diagnostic tests.

1.Randomized Control Trial (RCT)
2.Prospective (or cohort) Study
3.Retrospective (or case control) Study
4.Cross-sectional Study
5.Before/After Study
6.Other Common Descriptive Studies
7.Case Studies

For more detailed discussion, see references cited[2,3].

Evety study has a specific purpose. As there are numerous possible purposes there also exist numerous research methodologies. Designing a study is a creative process, but it is also entails professional skills (no different in this respect than the importance of a practising physician being clinically competent and current in the field). Selection of the components of methodology is not primarily a matter of sophistication, complexity or fashion, but depends on what is suitable to the purpose. Beyond the purpose, research questions and objectives must be carefully thought through and articulated, especially in relation to the underlying rationale and implications for study type and implementation. Both study type and implemtation require attention to operational considerations, which must also be incorporated within the protocol, such as: selection of study instruments, procedures, and variables; plan of Implementation details for data collection; and inferential plan for processing the data to develop generalizability of the findings.

Such steps are not as mechanical as they may sound. In fact they are as much art as they are science, by giving due consideration to the various implicit and explicit components social, moral, ethical and economical obligations. Even more fundamental, the clarity of wording a research question, reflecting a thought process, is pivotal to the focus and clarity of the intended research, with implications for study

type, implementation and analysis.
Moral: purpose determines design, and design guides the analysis

Operational Considerations

Designing a study starts with the broad framework just outlined, and operational details must then be developed. If this overall pattern is not followed (purpose and design first and foremost), one can get into an awkward situation. Imagine buying the buttons first and then trying to find a matching cloth for the formal dress of your dreams.

A research study design always progresses from macro to micro details.

The pertinent elements considered in any research design are:

1. Study populations and participants
2. Study variables, instruments and procedures
3. Time frames
4. Study setting
5. Logistics and budgeting
6. Communication

Study Population

A study's population refers to potential participants of the research, for whom the findings of the study are intended to be generalizable. The idea of the population in research is different from the common conception of the word, which is a group of individuals within certain geographical boundaries. In epidemiology, the population composition is based on the unit of analysis. It is not only limited to individual human beings, it could be households, groups, communities, even geographical regions. A researcher's role is to select the population to whom the research problem is not only relevant, but also accessible and feasible within the available means.

Sample

A sample is a subset of the study population comprising the actual participants of the research. The findings based on their data intend to be generalized to the study population. The researcher's role is to select the sample that best represents the population. The selection process of the sample from the population is called Sampling.

In order to assess the prevalence and risk factors of work related respiratory symptoms in Swiss farmers, a representative sample of 1542 Swiss farmers was selected[4].

When every member of the population has a probability of more than zero for being selected as a study participant it is called probability sampling. The most basic form of this is the random sample, or "equal probability" sample. First, all the members of the population are arranged in a list and serial numbers are assigned to them. Then the required sample size is randomly selected on the basis of those serial numbers. The whole process can be adequately done with the help of computer statistical software or random number tables. A random sample is considered more representative of the study population as compared to other less arbitrary methods. However, when sufficient resources are not available then non-probability sampling is sometimes opted. In this approach, not every member of the population may have a probability of being part of the study. Whatever method is used involves some subjective judgement in the criteria of selection (inclusion and exclusion criteria). Formal sample size calculations are required to determine the least number of participants required to make study findings generalizable to the population. Statisticians normally play a key role in guiding these aspects of study design.

Study variables, instruments and procedu res

In research, variables are either measured, controlled or manipulated. This requires attention to operational definitions, instrument validity, and other procedural details.

Operational definitions require first a decision on the type of data to collect.

Research was conducted to observe gender and age-related differences in corneal topography in the

normal population[5].
This study investigated the influence of aging and sex hormones on the cornea. In fact, the researchers main consideration was biological differences in relation to sex (which is biology), not gender (which is a role). While this may seem inconsequential to some, in reality it reflects the important issue of accuracy in scientific eommunications. An operational definition requires validation that what we are actually measuring is congruent with the theoretical concepts of the research problem as stated. Rejection of sex reassignment surgery among adolescents with gender identity disorder results in dysfunctional psychological profiles, as wa.c observed in follow up studies[6].
The Instrument is the tool used to collect the data (eg., questionnaires, scales, etc.). The instrument should have the capacity to capture the phenomenon in a manner that is valid, accurate and adequately precisely for the purpose intended.
A valid measurement is one that measures what it purports to measure. Accurate measurements are those, which are not contradictory to the true values of the measurements, while precision is the ability of the instrument to give the exact details of the measurement. An inaccurate measurement (faulty) could also be precise one.
Age of the study sample can be specified as less than or more than 40 years as younger and older adults respectively. This could he an accurate measurement but does not precisely specif$ the actual age difference between the study subjects.
Information regarding the actual age could be precise by asking the date of birth from the study subject. But there is a limitation since the measurement depends upon the memory of the study subjects and does not necessarily represent the true value.
Procedural details enable the researcher to explicitly delineate how the measurements are made. Such information makes measurements consistent and reliable. If anyone wants to replicate the measurement /experiment he/she can obtain the same results.
Moral: Tout par raison (everything according to reason)_____ Cardinal Richeiieu

Time element

There are a number of time specifications, which should be made clear in relation to the research design.

1. The time frame of the whole research exercise, from conception to the communication of the results.
2. The frequency and intervals between measurements should be specified.
3. The timing of measurements should be consistent in the research as circadian rhythm, daily routine and environmental factors (e.g., seasons of the year) may produce fluctuation in the value of some study variables.

Due to daytime fasting among Muslims during the month of Ramadan, changes occur in sleep, eating and social habits. This results in changes in the circadian regulatory mechanism of a number of body hormones such as Melotinin, FSH, Growth hormone, Cortisol and Testosterone[7].

4. The order of the measurements is also very important especially for those studies, which assess the temporal relationship between the variables.

Study settings

The selection of study setting in a study design is one of the first order / macro level decisions. Deciding on the study setting depends on the level of exploration and control allowed by the research purpose. Broadly there are three types of study settings:

1. Uncontrolled
2. Partially controlled
3. Highly controlled

An uncontrolled setting implies that the study is conducted in its natural environment, with no manipulation or control by the researcher. It is generally a better option for exploratory or descriptive

research.
Cognitive decline among the community dwelling elderly has a marked negative impact on their survival[8].
Partially controlled setting means conducting a study in an environment, which is controlled or manipulated to some extent by the researchers. It is generally not good practice to study the natural progression of a disease in a hospital setting, except perhaps in advanced or terminal phases. An exception to this is nosocomial infection, which by definition is a product of the hospital environment.
The incidence of Hepatitis C viral infection in renal transplant patients is up to 36% in India. Hemodialysis is the predominant source of this nosocomial infection and a majority of these patients remain asymptomatic during short -term follow-ups[9].
Field studies normally utilize either an uncontrolled or partially controlled settings. In field trials we manipulate certain variables with complete lack of control over the environment.
Field trials of BCG carried out in different countries have demonstrated that BCG vaccine is capable of preventing leprosi but with varied a protective efficacy of 20% to 80%[10].
Highly controlled setting means an environment, which is controlled or manipulated for all the independent, and intervening variables considered in the study.
Laboratory experiments are usually conducted in a highly controlled environment. Such settings are most favorable for the determination of causal relationships.
A study to evaluate the effects of dopamine antagonists concluded that they afftct the rate of accommodation of the human eye but not the degree of the accommodation[11].
The significance of the study setting decision is important not only for the feasibility of the study but also to determine its generalizability. Important differences in the findings of otherwise similar studies may be observed when treatment! manipulation is considered across highly controlled to partially or uncontrolled settings. Uncontrolled conditions are not always applicable and findings may be limited to the settings in which the actual study was conducted.

Plan of analysis

Statistical analysis is performed in order to transform crude data into manageable form from which the researcher may be able to develop new meanings. This step should be planned beforehand (a priori), and not after the data have been collected (a posteriori). An appropriate plan of data analysis presumes that the theoretical basis for the research question is known in advance; the alternative runs a risk of bias, for example selecting analyses only after having first looked at the data. In the extreme, an a posteriori search for meaning may produce chance findings (known critically in research circles as a "fishing expedition" or "data dredging"). Such an exercise may sometimes give rise to useful new hypotheses, but more often will produce spurious results. While this is not always a clear cut issue, and there is a legitimate role for exploration and hypothesis generation, the best professional practice is to determine the analytical approach, a priori and not to be influenced by extraneous considerations when the time for analysis arrives.

Factors to be considered in planning the data analysis include:

1. Research rationale, objectives, questions and hypothesis
2. Data quality, frequency and level of measurement
3. Study type
4. Resources, especially availability of computers and appropriate statistical software

The investigators' expertise in methods of analysis is sometimes critical. Just as the year of graduation is a predictor of clinical performance among too many practising physicians, so too there are many researchers who have not kept pace with new analytical techniques, or may be reluctant to refer to more knowledgeable colleagues in determining how best to approach a more challenging analytical issue.
For example, while statisticians can help in the analysis, medical researchers themselves must have enough understanding of statistical principles in order to be able to communicate their needs and

findings in a professional manner. Most statistical analysis is based on certain assumptions and investigator awareness of these is important for appropriate analysis. Beyond the requirements of statistical analysis, the researcher must also have the expertise to correctly interpret the statistical results in the context of the research. Careful syntheses of analytical findings with other biomedical and bio-social knowledge forms the basis of conclusion and recommendations.

Logistics and budgeting

This process is a significant component of research design and protocol development. It involves planning on how to best use available resources (always limited). It includes identifying all feasible options, and identifies responsible persons on the project team and how best to assign duties among them. Sound logistical planning and realistic budgeting reduces conflicts, especially among those involved in research and helps in timely completion of the study. From the standpoint of funding agencies of course, this process communicates how the required resources are intended to be utilized and is essential therefore in arriving at funding decisions. Effective research requires good management as much as it requires good science.

Corn m unication

Significant scientific findings in history become known not when they are unearthed but when they are properly communicated. The original observation of the ability of cowpox to prevent the occurrence of smallpox was made by farmer Benjamin Jesty some two decades before Edward Jenner was given formal credit for the discovery, following his presentation of a study on milkmaids to a scientific audience. Interestingly it took another two centuries before smallpox vaccination was universally adopted as a basis of the successful global eradication campaign. Communication of research findings, ultimately to the policy and decision making levels, therefore is fundamental component of research. However, it is mostly neglected in the planning phase. Nonetheless, one needs to plan for how and by whom research reports will be developed and decisions made regarding appropriate audiences and means to communicate with them. Conferences, seminars and electronic media are quick means to disseminate as compared to the publication of the findings in scientific journals. Choices can be made according to the urgency of the message and type of audiences. The internet has become a tool with enormous potential for dissemination of medical research findings and other medical information. Prompt communications is a key also to prevent the duplication of scientific work and speedy achievement of health care goals. The Human genome project is one success story of the innovative use of the Internet in which thousands of scientists from around the world participated by constantly sharing their findings[12]. From a users' perspective (general practitioners, students. researchers alike), such information is now much more quickly available on the internet when compared to more conventional methods of dissemination.

Moral: The simplest schoolboy is now /àmiiiar with facts for which Archimedes would have sacrificed his ljfe_____ Ernest Renan.

References

1.Sommer Ti. Suppression of scientific research bahramdipity and nulliple scientific discoveries. Sci. Eng. Ethics., 2001. 7:77-104.

2.Fowkcs FGR, Fulton PM. Critical appraisal of published research: introductory guidelines. Br. Med. J., 1991; 302:1136.1140.

3.Sackctt DL, Havnes RB. Tugwell P. Clinical epidemiology - A basic science for clinical medicine, Toronto, Little Brown & Company. 1985.

4.Danuser B, Weber C, Kunzli N, et at. Respiratory symptoms in Swiss fanners: An epiderniological study of risk factors Am. J Ind. Med.. 2001;39:410-18.

5.Goto 1, Klyce SD, Zheng X, et at. Gender- and age-related differences in comeal topography. Cornea,

2001 :20:270-76.
6.Smith YL, van Goozen SH, Cohen-Kettenis PT. Adolescents with gender
identity disorder who were accepted or rejected for sex reassignment surgery' a prospective follow-tip study J. Am. Acact Child. Adotesc. Psychiatry., 2001 ;40:472-81.
7.Bogdan A, Bouchareb B. Touitou Y. Ramadan fasting alters endocrine and neuroendocrine circadian patterns Meal—time as synchronizer in humans?, Life Sci., 2001, 68:1607-15.
8.Bassuk SS, Wypij D, Berkman LE. Cognitive impairment and mortality in the community-dwelling elderly. Am. ,J. Epidemiol 2000: 151 :676-88.
9.Agarwal SK. Dash SC. Irshad M. Hepatitis C virus infection during haemodialysis in India. J. Assoc. Physicians India.. 1999:47:1139-43.
10.Noordeen SK. Prophylaxis-scope and limitations. Lepr Rev 2000: 71 (Suppl:S.16-9: discussion) S 19-20.
11.Yeung L, 1 lung PT, Lin LL, ct al. Effects of dopamine antagonists in human eye accommodation. J. Ocul. Pharmacol Ther 2001:17:11-7.
12.Wolfsbcrg TG, McEntyre J, Schuler GD. Guide to the draft human genome. Nature, 2001:409:824-6.