INT J TUBERC LUNG DIS 8(8):1036–1040

Getting started in research: the research protocol

D. A. Enarson,* S. M. Kennedy,† D. L. Miller‡

* Department of Scientific Activities, International Union Against Tuberculosis and Lung Disease, Paris, France; † University of British Columbia, Vancouver, British Columbia, Canada; ‡ Emeritus Professor of Public Health Medicine, Imperial College, University of London, London, United Kingdom

SUMMARY

To be efficient and precise, research needs a 'road map', called the 'research protocol', which follows a standard format. It includes an abstract, study description, ethical considerations, significance of the study, the budget and a description of the investigators. Study description spells out the study question, the rationale for the study, including previous studies on the subject, the objectives, hypotheses and aims, design and methods, project management, strengths and limitations and a list of references. The objectives, hypotheses and aims are developed by outlining a general research topic (the objective), developing a hypothesis from the broad objective, translating it into the null hypothesis and then listing the steps by which the null hypothesis will be refuted or accepted (the aims). The design and methods describe the type of study to be undertaken, the population in which the study is carried out, including the sample size and statistical power, the selection of subjects, the methods of data collection, and outline of data management and statistical analysis. The detail of the protocol ensures that the study will be carried out successfully and is essential for all health research.

KEY WORDS: research; protocol; lung; education

IN LABORATORY RESEARCH, scientists generally employ an experimental design. Conclusions from following this approach are usually considered to be more 'scientifically sound' than those from simple, uncontrolled observations. In this paper we introduce the elements of a typical research protocol for population-based research. Despite the fact that most such studies are *observational* rather than experimental, it remains essential to elaborate, and follow, a rigorous research protocol. Doing so will increase the likelihood that the conclusions drawn from the research will be scientifically sound.

The first step before embarking on developing a detailed protocol is to select a research topic. In doing so we need to understand prevailing operational realities and how to work within their limits.

Example:
Points to consider when selecting a research topic:

- Relevance to health in the general community (frequency, seriousness)
- Likelihood of effective intervention or implementation of the findings
- Feasibility of successfully carrying out a study to answer the question
- Generalisability of the findings to other settings
- Interests and experience of the research team
- Relevance to political and social imperatives (community, scientific).

This article provides an overview of the research protocol. Subsequent articles will examine the individual components of the protocol in greater detail.

TYPICAL FORMAT AND ELEMENTS OF THE RESEARCH PROTOCOL

1 Abstract
2 Study description
  a Study question
  b Rationale, previous studies on the subject
  c Objectives, hypotheses and aims
  d Design and methods
    Study design
    Study population
    Sample size and statistical power
    Subjects: selection and definitions
    Data collection methods: measurement, definitions
    Data management and statistical analysis
  e Project management
    Personnel required
    Duration of the study (timeline)
    Follow-up procedures (if needed)
  f Strengths and limitations
  g References
3 Ethical considerations
4 Significance (or expected impact)
5 Budget
6 Investigators: role of each and curriculum vitae

Correspondence to: Professor Donald A Enarson, Director of Scientific Activities, International Union Against Tuberculosis and Lung Disease, 68 boulevard Saint Michel, Paris, France. Fax: (+33) 1 43 29 90 87. e-mail: union@iuatld.org

The research protocol can be totally fluid during its development phase but, once agreed, it must be strictly followed.

ABSTRACT

This should be concise but sufficient to orient the reader to the main purpose of the study, how it will be conducted and its expected benefits. It is a sketch plan of the study that will provide a view of the general plan before examining the details. *It is placed at the head of the protocol, but is often written after the protocol itself is completed.*

STUDY DESCRIPTION

The essential elements of a protocol as outlined above should explain the study in terms of answers to the following questions:

- *WHY?* Sets out the study question and the relevant background information
- *HOW?* Describes the study design and the rationale for choosing it
- *WHO?* Defines the target and study populations and sample size
- *WHAT?* Identifies the variables to be measured, instruments to use and outcomes to be analysed
- *SO WHAT?* Comments on the expected significance of results and contribution to knowledge.

Study question

The topic of the research is traditionally formulated into a question which in turn must be clearly defined in terms of specific objectives. The more precise the question, the more likely it is that research will provide new knowledge. By formulating the topic into a question, it is easier to define the steps necessary to arrive at an answer to the question. If the topic is stated but not precisely defined, it is less clear how the knowledge will be derived. The question format requires greater precision and leads to a logical approach to the research topic.

Example:

If the general topic of interest is asthma in Turkey, the research question could be: Why is asthma among children in Istanbul exceptionally frequent?

Rationale, previous studies on the subject

The purpose of this section is to state how the research question arose from current knowledge about the subject. The progression of your ideas needs to be set out in a logical sequence. Be concise; include key references, not a complete review of the literature.

- Discuss the importance of the topic
- Review the relevant literature and current knowledge (including deficiencies in knowledge that make the study worth doing)
- Describe any results you have already obtained in the area of the proposed study
- Indicate how the research question has emerged and fits logically with the above
- Outline your approach to address the research question
- Explain how your study will benefit the community.

Specific objectives, hypotheses and aims

Objectives

Even a precise study question is often too broad for one study to answer. Therefore, you must break down the question into one or more *objectives* for your particular study.

Example:

The objectives of this study are to determine if the excess asthma in Istanbul is due to genetic predisposition, socio-economic status and/or indoor air pollution.

Hypotheses

To meet these objectives, the study plan usually includes comparisons of disease or exposure rates between more than one group of subjects, using statistical testing to evaluate the comparisons. If this is the case, the objectives should also be stated in the form of *hypotheses* to be evaluated by the statistical tests. The hypotheses should be written as statements to be *refuted*. These are referred to as *null hypotheses*. The reason for this format is that the result provided by most statistical tests is the *probability of erring* if you reject the null hypothesis. Rejecting the null hypothesis increases our confidence, *with a given level of probability*, that there is a relation between the variables (health-related states, agents, determinants) studied.

Example:

The following null hypotheses will be tested:

1 Asthma prevalence rates are not different among children from low and high socio-economic groups in Istanbul.
2 Asthma prevalence rates are not increased in children living in homes with increased air pollution.
3 The relationships between asthma prevalence rate and socio-economic status and between asthma prevalence rate and indoor air

pollution do not differ according to the atopic status (as an indicator of genetic predisposition) of the child.

The progressive clarity achieved by formulating a topic into a question, a question into more precise objectives, and then formulating hypotheses should be obvious. The question sets the framework, the objectives force us to be precise about methods and to define key terms. Posing the research hypothesis forces us to think carefully about what comparisons will be needed to answer the research question, and establishes the format within which we will apply our statistical tests when interpreting our results. In determining the nature of an association, we can evaluate the strength of the association and estimate a level of confidence that the difference found in our study is true. These are two key elements in considering whether or not associations seen are likely to be causal rather than random.

Specific aims

Following this, you summarise the practical steps the study team will need to carry out to address the objective of your proposed study.

Example:

To meet these objectives, the following specific aims have been identified:

1 Identify a suitable source of childhood asthma cases and select 200 cases, following a specified case definition.
2 Identify and select suitable control subjects.
3 Record personal, demographic, and socio-economic information about cases and controls using a standard questionnaire.
4 Perform allergy skin tests on cases and controls (as an indicator of atopy).
5 Measure indoor particulate exposure on each of 3 randomly selected days for each participant.
6 Compare risk ratios for atopy, low socio-economic status, and increased indoor air pollution between cases and controls.

It should be obvious that you cannot complete this section of your specific aims immediately. You will have to decide on your study design and methods first, and then list your specific aims later.

Design and methods

Study design

This should state the selected design of the study. Explain why the particular study design has been chosen in preference to other possible designs.

Study population

Defining the group in which the study will be carried out provides the setting for which the research has relevance. This section also describes how one can be certain that the results can be generalised to the population identified.

Example:

- Will the results from our study population be able to tell us something about all children? About all children in Turkey?
- Who are 'children'? Those under 16 years of age? Under 2? Are there a lower and an upper limit?

Sample size and statistical power

It is necessary to estimate how big the study needs to be to answer the question posed. Specify the assumptions made for the calculation, and include a table of the calculation of sample size (and power) given varying assumptions.

Example:

- How many children will be recruited?
- Will this number of children be sufficient to answer the question?
- What is the variation in the prevalence of the factor (air pollution, atopy) among the children?
- How big a difference in children with/without the factor can I detect?
- Is this an 'important' difference?

Subjects: selection and definitions

This section should:

- explain how many subjects will be recruited, where and why
- define eligibility, inclusion and exclusion criteria, criteria for discontinuation
- estimate the numbers of potentially eligible subjects
- describe mechanisms of recruitment
- discuss the feasibility of recruiting the required number of subjects and estimate the proportion that will agree to participate.

Example:

- Will children with (without) asthma be included?
- Which children will be excluded?
- How will the children be approached?
- What will you do if they decline?

- How many children will be eligible and what percentage do you expect will actually participate?

Data collection methods: definitions and measurement

It is essential to state how the data will be collected to determine both the health outcomes (disease or other 'health-related state') and the determinants you are planning to study. Quality control procedures should also be specified. If the procedure is a standard one that has been described before, it should be referenced.

This includes:

- definitions of all terms
- pilot testing for methods and instruments
- the validity and reliability of the definitions proposed
- the limitations of the measurement tools and definitions proposed.

Example:

- How will asthma be determined for this study?
- What will happen if someone who doesn't have asthma is said to have it?
- What is the definition of indoor air pollution for this study? How will it be measured or estimated?
- How will socio-economic status be determined?
- What will be the definition of atopy for this study?

Data: management and statistical analysis

This section will describe:

- procedures for coding and entering data into computer files
- measures to ensure the completeness and accuracy of the information
- examples of how the results will be displayed and comparisons made
- tests to be used to carry out statistical analyses in order to test each of the stated hypotheses.

Include the appropriate reference for the statistical tests: the statistics book or article in which the method is described, or the statistical computer programme to be used.

Example:

- How will the patient records (or questionnaire responses) be abstracted?
- How will you ensure that the questionnaire data are correctly coded?
- What types of tables, graphs and figures will be used to display the results?
- What statistical tests will be used to test the hypotheses of the study?

Overall project management

This section will identify the *personnel* required to carry out the research and define their tasks. It will justify the personnel proposed in terms of the tasks and the amount of time required. It will specify the responsibilities of each staff member.

Duration or timeline will set out the anticipated time required for each phase of the study, including:

- pilot testing
- recruiting of subjects
- preparation of forms and questionnaires
- data collection
- follow-up procedures
- data checking and statistical analyses
- reporting—to participants, to sponsors, to the community involved, to the academic community.

Follow-up procedures for study participants should be specified where appropriate.

Strengths and limitations

In developing the protocol, many compromises will have to be made in choosing among several possible study designs and approaches to collecting information. It is important to include a section in which you address the possible criticisms of your design and methods and provide reasons why you think the limitations imposed by your choices are not serious ones. Similarly, it is useful to identify those aspects of the specific protocol that you think are particularly strong and worthy of financial support.

References

This section will list references from all the above sections, normally listed according to the order of their presentation in the protocol.

ETHICAL CONSIDERATIONS

This section must explicitly state that the principles of the Helsinki Declaration have been taken into account and will be followed. It indicates:

- how the quality of the technical aspects have been assured
- the expected hazards of the study procedures and their expected benefits
- the rationale and justification for carrying out the research

- the priority of the participants' interests over those of science or of society
- how these interests will be safeguarded
- responsibility for liability for injury to study participants
- how the participants are informed of the study, and
- how they give voluntary consent to participate.

If the research is sponsored from outside the community or the country, it must explicitly outline the interests of all parties in the research, who will be the authors of scientific publications, and the benefit to participants (including those who are diagnosed as ill), to the local investigators and to the community.

SIGNIFICANCE (EXPECTED IMPACT) AND KNOWLEDGE TRANSFER

This section restates the justification for the study in terms of the anticipated results. It will specify:

- the implications of the potential results
- how the results of the study may be used by your own research team in the future, by other researchers, by policy makers, by the community
- how the researchers will communicate the results to practitioners, policy makers, or community members who can translate the scientific findings into preventive measures, new practices, or treatments.

BUDGET

Each item of expenditure expected in the conduct of the study must be specified. A written budget justification should be included to explain various expenditures in further detail.

INVESTIGATORS: ROLE OF EACH AND CURRICULUM VITAE

This section should describe what role each investigator plays in the study, and should state clearly who is responsible for each component of the study.

The curricula vitae should provide a clear description of the qualifications and experience of the investigators, including training, academic degrees or certificates, research experience and scientific publications.

RÉSUMÉ

La recherche, pour être efficiente et précise, nécessite une feuille de route. Celle-ci s'appelle le protocole de recherche et respecte un schéma standardisé. Celui-ci comporte un résumé, la description de l'étude, les considérations éthiques, la signification de l'étude, le budget et une description des chercheurs. La description de l'étude détaille la question étudiée, la justification de l'étude, en incluant les études précédentes sur le sujet, les objectifs, hypothèses et buts, le schéma et les méthodes, la prise en charge des projets, les aspects positifs et les limitations ainsi qu'une liste de références. Les objectifs, hypothèses et buts sont développés en dessinant l'objet général de la recherche (l'objectif), en développant une hypothèse à partir de l'objectif général, en le traduisant en une hypothèse zéro et en faisant la liste des étapes par lesquelles l'hypothèse zéro pourra être soit rejetée, soit acceptée (les buts visés). Le schéma et les méthodes décrivent le type d'étude à entreprendre, la population où cette étude est menée, y compris la taille de l'échantillon et son pouvoir statistique, la sélection des sujets, les méthodes de recueil des données, la prise en charge des données et leur analyse statistique. Ce sont les détails du protocole qui assurent que l'étude soit bien menée avec succès ; il est essentiel dans toutes les recherches en matière de santé.

RESUMEN

La investigación necesita un 'mapa de ruta' para ser eficiente y precisa. Esto es lo que se llama 'protocolo de investigación', que sigue un formato estándar. Incluye un resumen, la descripción del estudio, las consideraciones éticas, el significado del estudio, el presupuesto y una descripción de los investigadores. La descripción del estudio detalla el problema estudiado, la justificación del estudio, incluyendo los estudios previos sobre el tema, los objetivos, las hipótesis y los propósitos, el diseño y los métodos, el manejo de los proyectos, los puntos fuertes y las limitaciones y una lista de referencias. Para desarrollar los objetivos, hipótesis y propósitos, se debe delinear la meta general de la investigación (el objetivo), formulando una hipótesis a partir de un objetivo amplio, el cual debe traducirse en una hipótesis nula y luego debe establecerse una lista de etapas, a través de las cuales la hipótesis nula será rechazada o aceptada (los propósitos). El diseño y los métodos describen el tipo de estudio que será emprendido, la población en la cual se realizará el estudio, incluyendo el tamaño de la muestra y su poder estadístico, la selección de los sujetos, los métodos de recolección de datos, el manejo de los datos y el análisis estadístico. La precisión del protocolo es lo que asegura el éxito del estudio y es esencial para la investigación en salud.