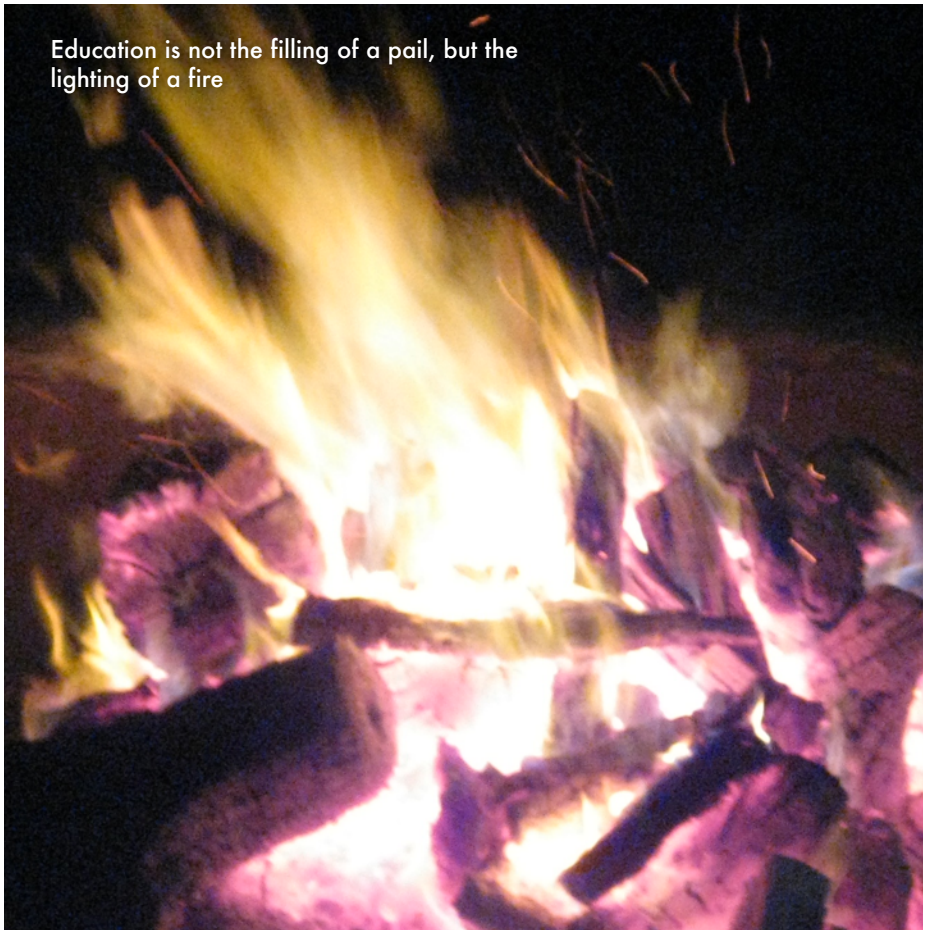
toolkit

WELCOME!



So, you want to do some research....

This is the start of a process that will transform you into a creator of knowledge.

Until now, your academic career has likely centered around memorizing the findings of others. By taking part in research you move yourself from being an outside observer, to a central participant in the process of shaping what we know and understand about the world. In medicine, achieving this transition will be set you apart as being a true scholar, and will positively influence the way you approach every aspect of your profession.

To begin, proceed to the general guidelines, which will help you come up with a topic and decide on a broad methodology. You will then make use of either document 3 (qualitative research guidelines) or 4 (quantitative research guidelines) in order to get detailed guidance for the type of project you have in mind.

Once you have completed your research, additional advice on publication and presentation can be found in the final section of the general guidelines. Good luck, and may your first experience as a researcher be a spark that ignites your academic imagination!

This toolkit contains:

1. Welcome document
2. General guidelines
3. Qualitative research guidelines
4. Quantitative research guidelines

Authors

Van Schalkwyk, GI; Botha, H; Bezuidenhout, J; Blitz, J; Van Schalkwyk, S and De Vries, J (UCT).

Advisors:

Van Heerden, B; Seedat, S.

This project was developed at Stellenbosch University Faculty of Medicine and Health Sciences, with support from the Centre for Health Professions Education and the Stellenbosch University Rural Medical Education Partnership Initiative (SURMEPI)

toolkit

DOCUMENT 2: GENERAL GUIDELINES

http://www.stockfreeimages.com/

Introduction

This is the first of three main documents comprising this research toolkit. This document provides guidance on issues that are applicable no matter what type of research you intend doing. They are therefore a good place to start your research journey.

Section one will help you to conceptualize your project to the point where you have decided on which broad methodological approach to follow. Thereafter, section two will help you formalise some of these ideas into a research protocol, which you will need in order to gain ethical approval for your research.

Section 3 provides some pointers on how you can go about presenting your research at a conference (either as a poster or oral presentation), as well as how to go about preparing it for submission to a journal.

This document should be used in conjunction with either document 2 or 3, depending on which type of research project you intend doing. Although every effort has been made to present things in a logical fashion, you will invariably find that you need to move between the documents from time to time. A useful approach is to get to the point where you have decided on a type of project, and then go straight to the relevant document for that project type. You will then be guided on when additional reference to these, the general guidelines, is likely to be of use.

Contents

Section summary

1. Introduction
2. Identifying a research problem
3. Crafting a research question
4. Choosing a methodology

Box 1: Why do research?

- Conducting research will help you to hone your own skills in critical analysis of existing literature. This will prove invaluable in helping you practice medicine in an evidence based fashion.
- Almost all forms of postgraduate study will require that you conduct some form of original research. This is often a stumbling block for students. Early practice will provide you with a distinct advantage.
- Conducting research can allow you to distinguish yourself from your peers in a quantifiable way, and can thereby assist the progress of your career.
- Our country has several major healthcare challenges. Equipping yourself with skills to conduct research can help you to develop socially responsive solutions to the problems that are most relevant to your local

1. Conceptualise your project

1. Introduction

The benefits of doing research as a student are numerous, whatever your career goals might be. Box 1 summarises some of these benefits. You may, however, be more concerned with other questions like – 'How on earth will I get time to do research?'. I mean, the typical day on a clinical rotation is not exactly what one would call relaxed. You wake up at 7h10, and decide to skip breakfast because you had to see your patients by 7h00. Once there you check on a few things, make sure the patient is still alive, have a quick listen to their lungs. Maybe this morning you will get to hear a coarse wheeze… or is it a crackle? Then it will be time for your ward round, and maybe you will get the opportunity to present your patient and learn something. Or you might just get a list of ward work and be told to go straight to theatre.

Theatre is no joke either. You stand in a corner trying to stay out of everybody's way, and gauge the progress of the surgery based on the particular scent in the air – burning flesh: they have just started. Something rotten: there's gunk in the abdomen. Lunch: they are closing, or at least they should be. So the doctor lets you take a break, and being the good student that you are, you decide to go and check on your patient before satisfying your stomach. You get to the ward and the nurse hasn't given your patient his morphine – again – and he is now curled in a ball of discomfort. Sigh. You confirm that it has in fact been written up, and take a moment to ponder the fact that the surgeon has started the patient on Amitriptyline. You hear Prof Niehaus' voice echoing in your head "Amitriptyline is NEVER the right answer!", but this is all a bit much now so you decide to sort things out after lunch.

At lunch you meet up with your friends on family medicine, who are done for the day, but are fiercely busy doing their patient write-ups. You take a minute to laugh at the case of a particular "Mrs Schoeman", who remains convinced that she is being poisoned by her anti-hypertensive's, and is insisting on a prescription of a particularly bitter form of green tea, mixed with an infusion of feline excrement. Ah, one has to appreciate the patients' perspective…

You cut your lunch short to go and do an incontinence rating scale that you were meant to do on a patient a few days ago. Lucky for you, it seems the patient has been discharged. Oh well, they can probably just do it at the clinic anyway. Besides, you have received an SMS to say the second ward round is starting, and today is not your firm's call day, so freedom is finally in sight!

2. Identifying a research problem

Perhaps your own experiences in hospital are a little different to those described above. However, it is likely that during the course of your studies you will at some stage experience situations similar to these. The reality is that, perhaps without realising it, you are on a daily basis exposed to multiple potential research questions. Let's take it from the beginning. Not all students wake up five minutes after they should be in hospital. Some, on the other hand, will be in hospital long before the sun has even thought of

coming up. And here is an interesting question – what motivates students to work? This may seem like an obvious question, but the situation of a student is rather unique. It is often thought that employees are motivated by receiving salaries. Students don't receive salaries. Students are often motivated by marks – but does anybody assess what time you arrive at hospital? The answer is not immediately forthcoming, and is a ripe topic for a qualitative research project. The next example comes as early as when you put your stethoscope over your patient's chest. You know what you are supposed to be hearing, but how do you know you are correct? Do you always identify sounds in the same way? Do you always identify them in the same way as your peers? Or consultants, for that matter?

3. Crafting a research question

Once you have decided on the problem, you need to try and frame it in terms of a research question. A quantitative question could be formulated by rephrasing the problem you identified about breath sounds as "What is the inter-observer variability of medical students assessment of breath sounds?", or another problem as "What is the likelihood of a patient having X-ray features of pneumonia if this is detected clinically by a medically student?". Qualitative questions about some of these problems could be phrased as simply as "What do students perceive as being the key learning opportunities during a routine ward round?", or "What factors influence the learning of students during time spent in a surgical theatre?" Still further qualitative questions can look at the perspectives of other members of the health care team, such as "Factors governing analgesic administration by nursing staff", or even patients "The views of patients on anti-hypertensives regarding the use of alternative medication". All of these questions have some important qualities – they are interesting, relevant to your daily environment, can be answered relatively easily, and have the potential to be published.

So this is then a good place to start. Think of questions that arise in your day to day experience of being a medical student. Decide on one that you think is interesting, but also one that you instinctively think you will be able to answer. This is something you may not know at the outset, but in the course of using these guidelines it is something we will help you to achieve. But first, you need to decide on something else – should your research be qualitative or quantitative?

4. Choosing a methodology

Table 1 summarises the main differences between these methodologies. Think about your research question, and try and see into which category it fits most intuitively. Also, think about the kind of research you have found interesting in the past, as that may influence what you would like to do for your own project. Finally, it may be helpful to do an informal literature review, to get an idea of how people have answered questions like yours (or, in fact, identical to yours) in the past. More detail on conducting a literature review can be found in the next section.

Table 1: A basic comparison of qualitative and quantitative methodologies

	qualitative	quantitative
Goal of research	Describe a phenomena	Quantify a phenomena
Typical methods	Interviews, focus groups, observation	Various techniques that generate numerical data
Data analysis	Data is coded and interpreted with the idea of describing main themes	Data is subjected to statistical analysis
Application	Ideal for describing answers to complex questions about what people believe and why	Ideal for analysing quantifiable phenomena, and determining relationships between variables
Sample project	"What do patients feel are important factors influencing ARV adherence?"	"What is the association between ARV non-adherence and CD4 count?"

Section summary

1. Introduction
2. Literature review
3. Describe your methods
4. Create a timeline
5. Prepare a budget
6. Discuss relevant ethical considerations
7. Highlight desired project outputs

Box 2: A literature review is…

- A summary and interpretation of key findings, theories and trends that exist in the literature that are relevant to your topic, rather than a simple description of previous research
- Where you cover recent research while providing a perspective of what has gone before.
- The place where you position your own study and provide a rationale for your research question.
- The place where you demonstrate how you will contribute to the on-going conversation in the journal.
- Where you define all the elements of your hypothesis, drawing on existing definitions and discussions.

2. Prepare a protocol

1. Introduction

Preparing a protocol is about more than just getting ethical approval for your project. A protocol will also prove invaluable in helping you plan and conduct a good quality research project. There are many ways one can go about preparing a protocol, but for the purposes of this toolkit, we will outline the method preferred by the Stellenbosch University Health Research Ethics Committee. Note that a full protocol is not the only requirement to have your project approved - see Box 3 for further requirements.

2. Literature review

As you commence with your research, one of your first activities should be to explore the literature to find out what others have done before. In other words you need to review existing scholarship. This will provide a basis for your own research and will give you ideas as to how you could go about your own study (see Box 2). In addition, the findings of your literature review will be one of the first things you address when you get to writing up your research for presentation and publication.

How does one start? This is often quite challenging for a novice researcher, but the availability of a multitude of online resources makes life a whole lot easier than it was for researchers twenty years ago! The university's library website is your most important port of call. This gives you access to just about every database (including Pubmed and Medline) that you will ever need. It also provides you with access to academic journals from across the world. An important component of starting your literature review has to do with delineating your research question and extracting key words (key concepts) from it. These become your search tools. Even if you start your search using Google Scholar, the most important choice will be the words that you choose to search with.

However, once you have started conducting online searches, you may be overwhelmed with all of the articles that your key words generate. You cannot possibly read everything, so you need to devise a working plan. Start with reading the most recent works. Read the articles that are cited most often (many databases provide you with this information). See which researchers seem to be mentioned most often in these articles – the most seminal authors. Find their work and read it, and then read the work of the people they cite. It is a bit like a scavenger hunt, with each new piece of information providing you with a clue to the next step. Of course, sometimes your search may come up blank. Note which words you key in when nothing of value seems to come up. This can be very exciting news because it could mean that there is very little research that has been done on this topic. However, this will also mean that you will need to read more widely and then construct a theoretical space for your own envisaged study. This will probably mean more work on your part, but it can be enormously rewarding to know that you are exploring a terrain that has not been explored before.

Initially, it is often a good idea to just scan the article to get a sense of what it is about. Read the abstract, the introduction (that is where the

author typically sets out her or his argument), scan the methodology section and then read the conclusion. If you think the article is relevant for your study, then go back to it and read it more carefully. Have a system for your reading, whether you are marking up on hard copy with different colours to indicate different themes, or whether you are using a software package and reading on line. When you read, have a checklist in your head: how have the authors constructed their argument, what literature (theory) do they draw on, what previous studies do they cite, which methods have they adopted, how did they collect their data, how did they analyse it and how is it presented, what are the key findings? This information provides important guidelines as you plan your own research.

Be consistent from day one, carefully store your sources and resources and spend time to set up a sound filing system. There are also many referencing programmes available, such as Mendeley, Zotero (which are free) or Refworks and Endnote which carry a cost. These programmes will save you many hours in the long run, so it makes sense to take some time at the start to set yourself up.

You can organise your literature review in a number of different ways – chronologically, according to the different theories, according to different groups of researchers, etc. Finally, plan your literature review carefully: what is your focus, which perspectives do you want to emphasize, what sort of coverage do you need to achieve, who is your audience? Does your literature review make sense given the answers to these questions?

3. Describe your methods

In this section you need to explain and carefully justify your choice of methodology. The way you do this will depend on whether it is a quantitative or qualitative project, and more details can be found in the specific guidelines for each of these project types.

4. Prepare a timeline

This should be done according to your preference, although the following check points are important:

- When do you begin data collection?
- When do you envisage being able to complete your data collection?
- How long do you think it will take to analyse the data?
- How much time have you allocated to write up your results?
- How much time have you allowed for editing and reviewing your work in conjunction with your supervisor?
- What is the endpoint for this project (paper, thesis, poster, etc.) and when do you intend to reach it?

Further details should be included as determined by the nature of the project. Additional guidelines are included in documents 3 and 4.

5. Prepare a budget

Whatever project you have in mind, you are likely to incur some expenses. It is important that you make every effort to anticipate what these expenses will be, and, in collaboration with your supervisor, ensure that you will be able to fund your project. The details of your budget are dependent on the nature of your project, but the following categories of expenses are important:

Box 3: Requirements for ethical approval

The precise requirements will vary, depending on which ethics committee is being applied to. The following is generally required:

- A sample of your consent form, which should be detailed yet accessible to your participants. Templates are often available from the relevant institutions.
- A full length copy of your protocol, with the following sections:
  - Introduction
  - Literature review
  - Methodology
  - Timeline
  - Budget
  - Ethical considerations
  - Desired outputs
- A summary of your protocol, with the following sections:
  - Introduction
  - Literature review
  - Methodology
  - Ethical considerations
- Declarations from all participants in the project, including supervisors.
- Permission from the relevant departmental heads whose departments may be affected by your research.
- In certain instances, fees may apply, although these are often waived in the case of undergraduate research.

- Printing costs
- Transcription costs (this is particularly relevant if you are doing a qualitative project)
- Transport costs
- Translation costs
- Professional costs (for example, if your project involves radiology, you may need to pay a radiologist to interpret your scans)
- Costs for statistical analysis

It is also important that you describe your sources of income, even if you will be funding the project yourself. If you intend applying for funding for your research, it is important that you do this well in advance. Requirements for funding will vary depending on the grant for which you are applying.

6. Discuss relevant ethical issues

There is perhaps a tendency for people to underestimate the ethical issues that may emerge during the course of their research. Guidelines can be obtained from the relevant ethics committee to which you are applying, and a variety of resources are available to guide your thinking in this regard. To begin, consider the following:

- Informed consent
  - What information will be included in your consent form?
  - How will you discuss this information with your research participant?
  - How will you explain to the participant that his decision to participate or not will not influence his medical care?
- Handling of information
  - Is it likely that your research may uncover sensitive, confidential information?
  - Will you have access to any sensitive information?
  - How will you ensure that the information is protected?
  - Will you be anonymising the data, and importantly, how?
  - How and where will you store the data?
- Publication
  - How will you ensure that your conclusions are accurate and reliable?
  - How will you describe ethical barriers that emerged during the course of your study?

7. Highlight desired project outputs

In the view of some scholars, and indeed, some ethics committees, it is unethical to do research that cannot lead to a meaningful outcome. Therefore, it is important that your protocol states exactly how your research will lead to a meaningful outcome. Essentially, this could mean either that your project will be:

- Submitted for publication
- Presented at a conference as an oral presentation
- Presented at a conference as a poster

If you are unsure about the eventual output, state the highest output for which you can reasonably aim. Remember, if you are uncertain whether your research is publishable, you probably need to revisit your research question and methodology before continuing.

Section summary

1. Introduction
2. Preparing a poster or oral presentation
3. Choosing a journal
4. Preparing a manuscript

Box 4: Key tips for multimedia

- Include key words and visuals – a picture is worth a thousand words.
- Include graphs rather than lists of numbers if possible.
- Do not include full sentences or lengthy texts – if you do want to include a lengthy quote, highlight its key words and phrases.
- Try to avoid reading from your slide unless you are doing it for specific emphasis.
- Do not use too many colours – develop a theme and remain within it.
- Stick to the same font (perhaps something different for headings) and use font size to indicate levels within the text.
- Never go smaller than point 20 font.

3. Publish your research

1. Introduction

Sharing your research with peers and researchers in your field is possibly the most important part of conducting research. If you do not share your findings with others, and expose your work to critical review, then it will be of very little value to your personal growth or to others. Ideally, therefore, you should always try to publish any research that you do. However, often it can be useful to first present your research at a conference or seminar within your own faculty or department, or at a conference which focuses on student-conducted research.

2. Preparing a poster or oral presentation

Once you have completed your research and have prepared your draft report/article as described in other sections in this Toolkit, you can start working on your presentation. Your decision as to what you hope to include in your presentation will depend on a number of different indicators such as the conference theme, who your audience will be (e.g. students, academics, teaching staff, clinicians, etc.), the format of the session (e.g. a poster session where you might have five minutes to introduce your poster), the duration of the session, what your "take home" message is and so forth.

For a poster it is important to consider how best you can present your study visually. If you are not artistically inclined, get some help. The best posters are those that are not over-full or text heavy, tell a visual story and capture the essence of the study. Typically conference posters vary in size and you should check the conference website for their particular requirements; it depends on the size of the boards on which the posters will be displayed. There are also many different materials and finishes available that provide you with options in terms of packaging and traveling.

If you are doing an oral presentation, then typically you would be expected to create a PPT slide presentation. Take time to do this properly rather than just running off a list of bullet points. Modern technology offers a wide range of options (video clips, online links, etc), but you need to apply your mind in terms of what will be the most suitable for getting your message across. A careful balance is necessary between tedium on the one hand, and too many bells and whistles on the other. If you are presenting in an unfamiliar venue, it is often a good idea to not include too many highly technical components into your presentation (you can't always be certain that the available software and hardware will be able to handle it). See Box 4 for some additional key tips.

A good presentation, like a good journal article, must be well-structured with a clear start, middle and end. Set clear goals for the presentation and revisit these at the end to make sure that you have achieved what you set out to do. Remember that research is often complex and you only have a limited amount of time to share your study with your audience. This means that you need to give enough information for them to make sense of your study, while at the same time ensure that you remain within your time limit. PPT offers many features that can help you to keep

the audience on track (e.g. running headers/footers). You may wish to create a visual map of the presentation that you insert as an icon on the slides at different points. This can act as a signpost and is particularly useful when you are presenting a strongly quantitative study with lots of data or a very complex piece of work.

Finally, practise your presentation for both timing and your ability to talk to the audience (rather than the computer screen or the projector screen). There are very few people who can 'wing it' in the academic world. Think about what sort of questions your audience might ask, and prepare insightful and well-considered responses. You put so much effort and time into conducting research, do not let a poorly prepared presentation detract from the quality and integrity of your work.

See Box 5 for guidance on how to apply for a conference presentation.

3. Choosing a journal

The first time you see your name in print can be a very exciting moment - getting there requires dedication and hard work! In an academic context, having a publication usually means that you have written a scholarly article that has been reproduced in a peer-reviewed journal. There are thousands of academic journals across the world and they cover every conceivable topic. It can often be quite overwhelming to have to select an appropriate journal, but the following might help:

- Remember that when you publish in a particular journal, you contribute to the scholarly conversation that is happening in that journal. When choosing a journal for your article, consider the extent to which what you have written can take this conversation forward.
- Have a look at your reference list. Are there perhaps one or two specific journals that were important and regular sources for your study? This might mean that your topic is in line with what the journal focuses on and, therefore, a good option for your study.
- Visit the journal websites – most often there is a description of what they aim to publish.
- Be realistic as to the quality and uniqueness of your work. Aiming for a high-impact, ISI accredited journal can be great, but it can be very disheartening to receive a flat rejection. Be strategic, know your strengths and acknowledge your limitations. Seek help from your supervisor and/or a critical friend. See Box 6 for more information on the concept of a journal's "impact factor".

Some journals encourage potential authors to submit an abstract of their work to the editor so that she/he can consider the suitability of the topic for their journal. This is often a good option to follow and can save you much time that would otherwise be spent preparing a manuscript for submission to an unsuitable journal.

3. Preparing a manuscript for submission

Once you have selected the journal, you need to start on your draft research report and turn it into a manuscript in article format for

Box 5:Presenting at a conference

The first step in this process is submitting an abstract to the conference for their consideration. An abstract is a short overview of what you hope to present, and will generally contain the same subheadings as your actual presentation. They will have a set of criteria according to which peer reviewers decide on whether your research is the type of research that they consider appropriate (content and quality) for their conference. It is best to follow their instructions for abstract submission very carefully – especially the number of words allowed. On the basis of your submission, they will offer you the kind of presentation that they think is suitable – poster, oral presentation, etc.

Box 6: Making sense of "impact factor"

"Impact factor" is a commonly used measure to evaluate and compare different journals. It is calculated by looking at the number of times articles in that journal have been cited in other article, and dividing that by the total number of articles published in the journal for a given time period. Hence, an impact factor of 3 means that on average, each article published in that journal was cited in three other articles. Although impact factor is a commonly used statistic, it is a controversial measure, and should not be taken too seriously. Rather consider what your research says, and who the most appropriate audience will be, and aim for a journal that will allow you to reach them.

submission. Carefully read through the journal's instructions to authors to ensure that you meet all of the criteria. Important issues are number of words, formatting, and citation (commonly Vancouver) conventions. For example, if your draft report is already 6000 words in length, but the journal requires only 3500, you will need to decide what to cut and what to focus on. Scan some of the articles from the journal to help you in this. What is the norm? Do they seem to prefer a short introduction, but a detailed methodology section? How do they present their data? Do they integrate their findings and their discussion? What emphasis is placed on making recommendations? Is there a restriction on the number of tables or figures? Do they require keywords? And so forth. If your study was quite complex or produced multiple themes, you may need to decide to only focus on one aspect of the research and then keep the rest for a second article. While you are scanning, look at the style, that is, how the different authors write. Consider whether or not the language is formal and academic. Are the articles written in the passive voice, or do they use the first person (less common in quantitative work)? To what extent can you see the author's voice emerging from the text?

When you submit an article for publication in an academic journal it will usually go through a peer-review process. Although this process can differ slightly from one journal to the next, it generally means that the editor will look at the topic and then assign the article to two or sometime three fellow researchers who are experts in that particular field. Most review practices are anonymous. Participating in this process can be very fulfilling, especially for a novice researcher. Journals generally draw on an international body of reviewers, so your article could reach many corners of the globe. While not all reviewers are as dedicated as they might be (and you must steel yourself for harsh rejections sometimes), many are excellent and you will learn much about your ability as a researcher from these reviews. Keep in mind that the first time you submit a manuscript for review, you join a unique community of practice where academics the world over spend their precious time on reading, considering and then offering feedback to their colleagues – anonymously and for no tangible reward or remuneration. Be sure to one day take up any opportunities you may get to review the work of others to ensure that this community lives on!

Publication is a slow process. It will take a few weeks for the editor to consider your manuscript and assign it to suitable reviewers. Reviewers are usually given about six weeks, then their responses are collated by the editor and sent to you with a decision – accept without revisions; accept with minor revisions; accept with major revisions, or reject. If you have been asked to revise, you usually have about a month. Once approved, the manuscript will be logged for publication, but this process can equally take several months – that is why you will often find references given as 'in press' or 'forthcoming'. However, online publications are becoming increasingly popular and speed up the process significantly – another option for you to consider.

Ultimately, the manuscript that you submit for peer review should be flawless. Cross-check references, ensure that you have not plagiarised anything, and check numbering, spacing, cross-referencing and so forth. Be as meticulous as you can, proofread your work carefully and ask someone else to read it as well. Your first publication could be your first step towards a rich and rewarding academic career – good luck!!

toolkit

DOCUMENT 3: QUALITATIVE GUIDELINES

Introduction

There are some questions that numbers just cannot answer.

This may be obvious for questions like - what is the meaning of life? What does it take to be a good doctor? These are questions that have no definite answer, and about which people may continue to argue. But then there are also examples that are a little less obvious. Lets take the example of a high prevalence disease in South Africa, like HIV. Why do some people default on their treatment? This is certainly an important question, and one worthy of answering. But how do we actually answer this?

One way would be with a quantitative study. We could do something like a questionnaire, and ask people to choose from a list of possible answers, which we then count up and draw conclusions from. This is an option, and certainly a commonly used method, but it is not guaranteed to give us all the answers. For instance, there may be many factors influencing drug adherence, and we would need to understand all of them before we can say something sensible about how best to address this problem. Qualitative research is less focused on counting responses, but more about trying to understand a problem in its full complexity, by exploring the *how* *and why* of a specific response. This approach is becoming increasingly common in medicine, as people begin to realise that the problems we face are of such a complex nature that they cannot be solved in the laboratory alone. We need to pay equal attention to how people think and behave, and consider how this may influence our approach to medicine.

contents:

Section summary:

1. Decide on a research question
2. Do a literature review
3. Choose a data-collection method
4. Decide on a sample
5. Make a timeline
6. Draft a protocol
7. Submit to a chosen supervisor
8. Submit for ethical approval

Box 1: Choosing a supervisor

Choosing a supervisor is a balance between approaching someone you feel comfortable with, and someone who is in expert in that field. Bear in mind, if you have a willing supervisor with limited expertise, you can always get further input from someone else. A reluctant supervisor with a lot of knowledge may in the end be more of a challenge. Try and ensure that you get your supervisor involved as early as possible. Ensure that you are on the same page about what you hope to achieve with your research, and how you intend going about it.

1. Plan your study

1. Decide on a research question

Deciding on a research question is the most important part of planning your project. In formulating your research question you will need to take the problem or issue you are interested in, and consider the types of related questions that can be answered using qualitative research methods. Let us continue our example. We might know that we are interested in why people do not take their ARVs. But to simply ask "Why don't people take their ARVs?" leaves us with a difficult task indeed, and one which would probably require years of work and could lead to a PhD. Perhaps we could narrow it down to say - "What do patients cite as the most important factors in ARV adherence?". By focusing on patient opinions, we might not answer the ultimate question, but we can certainly add value to the discussion. Remember, a humble and well-executed project is better than one that is ambitious, yet poorly conducted. Our new question could be even further refined by defining the parameters of the group we intend to study - for example, we could rephrase it as "What do patients at a primary care ARV clinic in the Western Cape cite as important factors in ARV adherence?". Write down your own research question in the space below. Make sure it is in pencil - you may need to change it later.

_____

_____

2. Conduct a literature review

In this step, you look at what others have said about topics relating to your research question. The literature review is a very important component of your research project, as it will influence how you go about your study as well as how you interpret your results. Furthermore, a brief literature review is necessary in order to submit a protocol to an ethics committee. While working on the literature review you will need to locate relevant papers and books on the topic you have chosen, read and summarise these papers and describe how they relate to your research question. Finally identify gaps in our knowledge - those things we do not yet know that we should know in order to address the problem. A detailed guide on how to conduct a literature review can be found in Document 2.

3. Choose a data-collection method

In qualitative research, three of the most commonly used ways of collecting data are semi-structured interviews, focus groups, and observational methods. Table 1 outlines some of the advantages and disadvantages of each of these methods. Choosing which method is best suited to your study involves several considerations. Observational methods are best suited to answering questions about how people behave within a certain context. Such studies can be potentially powerful, but for the purposes of these guidelines we will be focusing on the other two methods – interviews and focus groups.

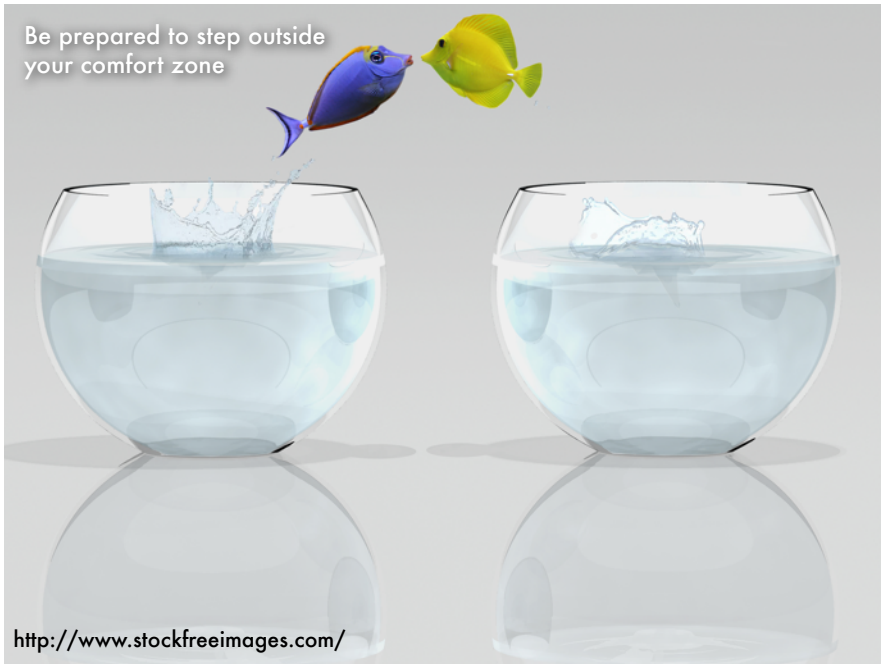
Table 1: Choose a method	observation	interview	focus group
Observational studies, semi-structured interviews, and focus groups are the most commonly used methods for collecting qualitative data. You may decide to use a combination of methods for your study.	**When?** This method is useful for describing how people behave within a specific environment - for example, how do patients behave at an ARV clinic?	**When?** With interviews you can gain a detailed understanding of the views of an individual. The intimate nature of the process allows for discussing sensitive topics.	**When?** A focus group helps you gain a broad range of perspectives within a relatively small space of time. Focus groups are often used in conjunction with interviews.
	Pros: This technique, if properly conducted, can give you the most accurate idea of how people truly behave within a certain context.	**Pros:** This technique allows for the best chance of gaining a thorough and detailed understanding of a specific view.	**Pros:** This technique can provide a large amount of data in a short space of time, and can allow for a broad range of perspectives.
	Cons: This technique may be technically difficult to execute, and can be very consuming.	**Cons:** You may miss out on important themes within the population if you do not conduct sufficient interviews.	**Cons:** Managing a group can be difficult. Discussing sensitive topics within a focus group may be inappropriate.

A semi-structured interview allows you to obtain the in-depth opinions of individuals about a specific issue, and relies upon the development of an interview schedule (see Table 2). What makes the interview "semi-structured" is that although you have a list of issues or questions you wish to explore, you allow yourself the liberty to be guided by what your participants are saying. A good interview schedule will either be a list of open ended questions or a set of more general prompts on topics that you would like to cover in the interview. Both questions and prompts are organized around the themes that you identified in your literature review, or in discussions with your supervisor.

The value of a semi-structured interview is that unlike a simple survey, in which you are restricted to sequential responses to all your questions, in a semi-structured interview you may choose to ask additional questions and modify your interview based on the participants' responses.

Focus groups can be used to answer similar questions to semi-structured interviews, although the data you obtain will be different. Whereas a semi-structured interview allows you to explore an individual's views in great detail, a focus group creates a dynamic environment in which you can gain understanding of how an issue is understood within a group. Generally the goal is to get a broad range of perspectives about an issue in a relatively short space of time. It can be difficult to conduct such a group, although a variety of strategies can help in establishing a more positive dynamic. Focus groups are seldom appropriate for discussing very sensitive or personal issues, and in such instances semi-structured interviews may be more appropriate.

Choosing between these methods can be difficult, and aside from the type of data you hope to obtain, the choice can also depend on the resources available to you. Think about how much time you have and think about how you will feel more comfortable – it is no



use picking the method that is theoretically ideal if you won't feel comfortable discussing the issue in that way. An alternative is to use a combination of both techniques, to gain an even broader range of data.

Whatever technique you choose, you will almost certainly be recording your interactions using an electronic recording device. This will then allow you to make a transcription of the encounter, which will in conjunction with your field notes (notes that you make throughout the research process), serve as the basis of your data analysis. You need to make sure that you include this in your methods sections, as well as when discussing ethical concerns relating to your project.

4. Decide on a sample

Two factors are important here. Firstly, you need to decide who will be the target of your study. When there are multiple stakeholders in a specific issue - for example, in the case of ARV adherence, you need to decide which perspective you are most interested in obtaining. See Box 4 for some ideas. You can look at what previous authors have done, and either do your study in a similar type of sample, or choose to explore a different perspective, in which case you can draw attention to specific contrasts in opinion.

Once you have decided on the group on which you intend focusing, you need to decide how you intend recruiting participants - this is known as a sampling strategy. *Random sampling* is a self-explanatory method that is best when you are hoping to gain a representative view from a relatively homogenous group of people. For example, if the focus of your attention was attendees at a specific ARV clinic, this would be a good strategy to employ. Make sure that your sampling is truly random - preferably by randomising an electronic list of potential participants. Picking patients "at random" whenever you happen to be able to go to the site where you are recruiting is not random sampling, but rather what is known as *convenience sampling*. Although this technique can be used, and in certain cases will be your only option, it is less rigorous than a true random sample.

Purposive sampling is a strategy where you actively seek out the input of specific types of people within the population you are studying. For example, you may wish to ensure that you have at least one person from each of the most common racial, language, cultural and religious groups in the community. This is particularly valuable in instances where the population under study is more heterogenous.

Box 2: Example timeline

Project - student learning in theatre

1 January
Start planning project, begin working on draft protocol

14 January
Submit draft protocol to supervisor

14 - 31 January
Finalise protocol in conjunction with supervisor

1 February
Complete ethics submission

21 February - 21 March
Gather data

21 March - 28 March
Conduct preliminary analysis

1 April
Discuss possible findings with supervisor. Decide on what you think can be reported

2 - 14 April
Do formal data analysis

15 - 30 April
Write up results

1 May
Submit draft article to supervisor

1 May - 14 May
Finalise article in conjunction with supervisor

15 May
Submit article to suitable

Box 3: Sample qualitative protocol: methods section

This study will take the form of a qualitative research project. This technique will allow us to gain a more detailed understanding of participant's views. To this end, we will develop a semi-structured interview schedule, drawing on themes that have been identified in the literature as being of greatest importance. Thereafter, we will be conducting a focus group discussion with nurses at an ARV clinic, to gain a contextually sensitive perspective on our interview schedule. The focus group will be conducted by the primary investigator, with annotations made by one of the co-investigators. The results of the focus group will be used to modify the interview schedule as needed. We will then seek to enrol 20 to 30 participants to participate in semi-structured interviews. Our sampling method will consist of convenience sampling. We will approach participants attending the Bothasig ARV clinic on specific days, and ask them whether they would be willing to participate. Written, informed consent will be obtained from willing participants. Interviews will then be scheduled according to the preferences of the patient. Interviews will be conducted in quiet, private spaces. The interviews will be recorded and transcribed. Data will be analysed inductively, allowing themes to emerge organically. The main themes will then be described, whereafter the significance of our findings will be discussed with appropriate reference to existing literature. Where possible, appropriate recommendations for further research will be indicated.

5. Develop a timeline

This can be done according to your preference. You should allow at least six months to get a qualitative project from start to finish. See Box 2: Example timeline.

6. Draft a protocol

In order to conduct your study, it is necessary to write a protocol, which will be submitted to the university ethics committee for review. The precise steps required to complete a protocol for the ethics committee can be found in Document 2.

When you describe your methods, do so as simply as possible, describing what you intend to do, and why you chose that method. Adding a reference to an existing study that used this method can add legitimacy to this component. This section must include a description of how participants will be enrolled, how data will be collected, and finally how data will be analysed. See Box 3 for an example of how this section can be structured.

7. Submit to supervisor

It is best to make contact with a potential supervisor as early as possible to make sure that they are available and to get their input on your research question. Depending on you and your supervisor's preference, they may wish to comment on your draft protocol at various stages, but it should occur no later than at this point. See Box 1 for advice on choosing a supervisor.

You will need to give your supervisor enough time to read through your draft. He or she is likely to have several comments and/or changes. Accept these with good grace - one day you may be in their position. For now, try and learn as much as you can, and rely on their experience. Doing things as your supervisor suggests at this early stage will also make them more likely to fully support you for the duration of your project.

8. Submit for ethical approval

Once you and your supervisor are both happy with the protocol, it can be submitted to the ethics committee for review. A detailed explanation of this process can be found in Document 2. Template documents may be available from the website of your institutions ethics committee.

Box 4: Populations to study

- Patients - the most obvious one. Medicine exists to serve patients, assessing their views on all matters is essential
- Doctors - as the core of almost any healthcare system, interviewing doctors is likely to yield important and well-considered opinions
- Nurses and other healthcare professionals - interviewing other members of the healthcare team can reveal interesting, previously unconsidered perspectives.
- Family members and caregivers - speaking to caregivers can either serve to enhance your data, or form the basis of a study if the research question is suitable.

Section summary

1. Make an interview/ discussion schedule

2. Do a pilot interview

3a Conduct interviews

3b Conduct focus groups

4. Get data in correct format for analysis

5. Analyse your data

2. Do your study

1. Develop an interview/discussion schedule

The interview or discussion schedule is the document that will guide you during the process of conducting interviews. It can be based on a number of sources. Firstly, there is your own personal experience. For instance, in our example of ARV compliance, you can probably identify some factors influencing treatment adherence that you have encountered in your clinical experience. Secondly, you will draw on your literature review to develop your schedule. Most importantly, you will draw on the themes others have found to be of relevance to the issue you wish to explore. You may also want to have a look at interview schedules that others have used, for instance other researchers at your university, or ones available online.

Once you have decided what themes you wish to explore, you need to phrase them into a short series of open-ended questions or prompts (or a combination of both). Here there is some room for individual preference – some people prefer to work with a list of the themes they wish to explore. Others like to have a list of specific questions grouped under particular themes. You will need to find out what works best for you. Whatever you decide, however, you will need to ensure that your questions or prompts are organised logically around particular themes. For example, if you have noticed that patients tend to misunderstand the importance of taking ARVs on daily basis, you could ask "How much do you know about taking your ARVs?". In this way, you start with a very open ended question, and you can then explore this more fully, either by picking up on participants' responses, or by carefully prompting them with phrases like "And is it important to take them every day?".

Deciding how many questions and prompts are needed is difficult to do in advance, and depends greatly on how much your participants know about what you are asking and the extent to which you are able to draw additional information. This is part of the flexibility that makes qualitative research such an interesting method. See table 2 for a sample interview schedule.

2. Do a pilot interview

Once you have developed your interview schedule, it is important to do a couple of pilot interviews with a friend, colleague or even a willing patient. This will serve several purposes. Firstly, it will help you see whether the questions you are asking lead you to discuss the issues you want to explore. This can be difficult, as the dry run is likely to be done with a peer who may not respond in the same way as a member of the population you wish to study. Still, you can gain some sense of whether others understand the meaning of your questions in the same way you do. Also, you can get some idea of the logical flow of your interview – you may find that questions you ask at the end would be better addressed earlier on in the interview. You can also determine whether or not your interview schedule appears to have either too many or too few questions. Finally, you grant yourself the opportunity to practice the interview schedule, so that you will be familiar with it when you actually start the process of interviewing participants.

Table 2: Interview schedule: ARV compliance

Questions	Prompts
1. Tell me what you know about ARVs?	Names, dosages, side-effects, appearance of tablets, resistance, etc.
2. Do you think it is important to take ARVs regularly?	Why?, who told you, under what circumstances, etc.
3. What do you think could happen if you don't take your ARVs as prescribed?	Relapse, resistance, viral load, CD4 count, etc.
4. Do you ever struggle to take your ARVs as prescribed?	Why?, when?, how often, etc.
5. What other reasons do you think make people not take their ARVs as prescribed?	Lack of understanding, pill burden, side effects, access, etc.
6. What do you think could be done to make it better?	Smaller tablets, safer medications, better education, delivery at home, etc.
7. Is there anything you would like to add?	Here, you can raise anything you felt was interesting that you would like to hear more about.

Other ways to "test" your interview schedule include conducting a focus group, or sending it to experts for comment. The information you will get from doing either of these will more likely relate to whether or not you are addressing the issues that people feel are important, rather than the technicalities of how your schedule performs in a true interview situation. Once you have done your pilot, you can modify your interview schedule based on the feedback you have received, and your own experience of how it performed. This process is very important both for ensuring that you have the best interview schedule possible, but will also be important when you get to the point of describing how you collected your data. The more thorough you are at this stage, the easier it will be to convince readers that you conducted high quality research.

3a. Conduct interviews

First we need to consider some practical aspects. The time and place of your interview is something you will need to discuss with your participant. This will vary depending on the nature of your project and the sampling method you have employed. Common sense will prevail here, but some general considerations are worth mentioning. Ensure you are somewhere where interruptions are unlikely, given the fact that you will be recording the interactions. Transcription will be more challenging if other voices can be heard intermittently. Excessive ambient noise can also be a problem. At the same time, it would be better if your participant is in a comfortable situation, as this will improve the quality of his or her responses. Interviewing participants in their homes is a possible solution, as long as a sufficiently private space is available. Remember that wherever you decide to conduct the interviews, you will need to ensure that there is some private space available, especially when you are investigating sensitive research questions.

Once you have met up with the participant and introduced yourself, it is important to obtain informed consent. This process is more fully discussed in Document 2. It is particularly important in qualitative

research that you explain to the participant that you will be recording the interaction and why. Once everything has been explained, you can provide the participant with two copies of your informed consent document. The first copy they may retain as a reference, the second they may review and then sign if they are satisfied.

It is important that the process of taking consent be viewed as an opportunity for building trust and rapport, rather than a barrier to completing the interview. If your informed consent process is appropriately participant-centered, and you take care to explain to the participant exactly what you intend to do and why, you will feel more satisfied with the process and will help set potential participants at ease.

Now for the actual interview. Make sure you have the interview schedule in a position where it is clearly visible. Also ensure that you are able to take notes, preferably on spaces left between questions on the printed schedule. You can start the recording, placing the recording device somewhere between yourself and the patient (you will need to test the device beforehand to ensure that this will lead to adequate sound-quality for transcription). Introduce the interview loudly and clearly, stating the date on which it took place, as well as the number of the interview (eg "This is the 4th of September, 2011, and today we are interviewing participant number 6"). Then collect necessary demographic information, after which you can proceed with your actual questions.It is important to adapt the interview throughout your project, based on the responses you are getting. You may for instance find that some of the themes are more important than others, and may want to focus more attention on understanding those themes during the interview. Or you may find that the way in which you mapped out the interview does not work as well as you hoped.

Box 5 summarises some of the main communicative strategies which you can employ to more fully explore each response. Make liberal use of these techniques. Try and stick to open ended questioning as far as possible, and avoid asking a leading question. If there is something you really would like a participant's opinion on, consider other ways of getting it other than asking leading questions. For example, you can try pick up on something they may have said earlier. Carrying on with our example of ARV compliance: say you want to know whether the clinic sisters are, in the views of the participant, taking enough time explaining the concept of viral resistance. You could say "Do the sisters tell you enough about viral resistance?". Doing so would likely give you a 'yes' or 'no' answer, but would give you data of a quality similar to a basic survey. In qualitative research, one can do better. Perhaps when you asked an earlier question, like "How much do you know about ARVs?", the participant said something like "... AIDS becomes resistant or something...". You could then pick up on this at a later stage by saying: "You mentioned earlier how AIDS becomes resistant. Tell me a bit more about that?". You could then prompt them by asking "who explained that to you?". Here they may refer to the doctor, or the nurse, or another source. At this point, it is safe to then ask "what did you think about the nurse's involvement in all of this?". You end up getting to a similar question, but by guiding the participant through a careful process based on his or her own responses, the end result is likely to be a well considered answer, richly contextualised by the participants themselves. The quality of this data is likely to be superior.

During the course of the project, it is vital to take field notes, recording both verbal and non-verbal responses during the interview, as well as your thoughts about how your interviews are helping to answer your research questions. Depending on your preference, make notes either during or soon after conducting each interview. You may wish to specifically note any quotes you think are illustrative of a specific viewpoint. Remember, although you will later have access to all your recorded data, you may miss things during your analysis, and having good field notes can both reduce this risk, as well as make the process of analysis that much easier.

Interviews can be of variable length, depending on many factors. Most often you will know the interview is completed - either because you have exhausted your questions or because participants become disinterested.

Once the interview is complete, thank the participant, and give them the opportunity to ask questions. Then stop the recording and thank the participant again before you leave. It is important to find a place where you can make some additional notes as soon as possible, reflecting on what stood out for you in the interview. Also reflect on things you may wish to do differently in the next interview, emerging patterns, new questions and the like. If it is still early on in the course of your study, you may choose to change your interview schedule. But even in your last interview, you may pick up on new ways to encourage participants towards giving good responses. One of these would be to take findings from previous interviews, and test them with subsequent participants.

Box 5: Helpful communicative strategies

- The simple probing question: Never underestimate the value of simply saying "can you tell me a bit more about that?". If that fails, try be more specific, saying "can you tell me how that affected you personally?"

- Reflective summaries: This has a two fold purpose - by presenting the patient with your own summary of their responses, you give them the opportunity to correct any false assumptions, and provide additional detail.

- Linking new questions to past responses: Often, when answering one question, participants will inadvertently answer part of another. When you get to that question, you may get better responses by referencing this prior response, for example you could say: "Previously you mentioned that the pills are very big. Are there other things about the pills that bug you?

- Assume a position of ignorance: Participants may be hesitant to volunteer information that they think you already know. This may be the case, but it is important to get their understanding. Ask the participants to respond as they would if they were talking to someone with no understanding of the topic under question.

For example, you could say "A lot of people have told me that ARVs are too big to swallow. How do you feel about that?".

3b. Conduct a focus group

Many of the requirements for planning for an interview are also relevant for a focus group. Logistically, however, it is often more complicated organising a focus group session and you will need to prepare well in advance. Your venue needs to be big enough for everyone to sit comfortably, preferably in a circle or around a table. This facilitates communication and it also makes for easier recording. Usually you would aim to have between 5 and 8 participants and it will be important that you carefully check your recording equipment before the interview.

One of the key challenges with focus group interviews is actually getting participants to attend the sessions and to arrive on time. This may require quite a lot of pre-interview activity on your side such as follow up emails and text messages to remind people of where they have to be and at what time. It is often a good idea to invite the maximum number of participants to cover for any 'no-shows'.

The most important aspect of a successful focus group is that it encourages easy and open communication. Often having some snacks and refreshments can facilitate creating a comfortable atmosphere. Just make sure that everyone has poured and dished before the recording starts! Consider your role as interviewer carefully before the interview starts. What is your relationship with the participants and how might you influence the atmosphere in the room? For example, if you are interviewing people from the community for the ARV research project mentioned earlier, consider how they are going to feel about a young person asking them about how they take their medicine? To address this you probably need to present yourself as professionally as possible without being too formal. Consider your body language – is it relaxed, but still demonstrating a clear interest in what is being said? Watch out for fidgeting, or looking bored. Consider your facial expressions so that you don't show the group when you disagree with what they said or think their response is inappropriate. As with the individual interview, be sure to explain the purpose of the research and obtain everyone's consent before starting the discussion. If the group members do not know one another, encourage them to introduce themselves at the start as you switch on the tape recorder.

During the interview itself, it is your responsibility to keep the conversation going without actually contributing to it. Do not insert your own ideas, rather ask others in the group what they think and how they feel about what a previous speaker has said. Watch out for participants who do not contribute and incorporate strategies to draw them into the conversation such as going around the table to give everyone a chance to comment on a particular issue, and so forth.

Another responsibility that rests with the facilitator is that of managing the time. A good time span is something between 60 – 90 minutes although this will depend on the size of your group and the topic under discussion. Warn your participants up front about how long you expect to keep them, and stick to that. If you feel that too much time is being spent on a particular discussion, intervene and move on. For example: 'this is

clearly a topic that people have many ideas about, but why don't we move on and see what else needs to be considered ...'. If you run out of time, but have the sense that people still have important contributions to make that will be of relevance for the research you may either need to reschedule, or offer participants the opportunity to submit their thoughts in writing.

Finally, often one can learn a lot about how people feel by just looking at their body language. It can, however, be difficult for you to manage the interview (making sure that you are picking up on all the topics in your interview schedule) and watch how people behave during the interview. For this reason it is a good idea to have someone else sit in on the interview as a scribe. That person can take notes during the session similar to what we described earlier when discussing note taking during the individual interviews.

4. Get data in correct format for analysis

The primary task here will be the transcription of your recorded interviews or focus groups sessions. These transcriptions, together with your field notes, constitute the data you will be analysing. Transcription is not a technically difficult process, but can be very time consuming. As a general guide, ten minutes of recorded interview will take about an hour to transcribe. Therefore, if possible, see if you can locate any funding from your supervisor or department to pay for a professional transcriber..

Even if someone else is doing it for you, transcription can take a long time, so it is best to get the process going as soon as possible. If you have finished your data collection and are still waiting to receive your transcriptions back, you can take this time to read through your field notes. Try and group the ideas you have noted into organised categories. At this stage, however, don't discard anything - you will only be able to know what is important and what is not when you have started analysing your transcripts.

5. Analyse your data

The process that you will now undergo is called coding. You can begin this process as soon as you receive the first completed transcript. There are two main approaches - deductive or inductive analysis. Deductive analysis is where you look at your data and try to identify references to a predetermined set of themes. Look at the included extract (Table 3), which is a fictitious transcription following an interview about ARV compliance. If we were to employ a deductive approach, we would read the extract, and look for any references to themes that we have previously identified as being relevant, usually based on what we have found in the literature. As we saw in the example literature review for this project, we

Table 3: Sample transcript extract

Inductive analysis (what emerges from data?)		Deductive analysis (using themes from literature)
"really big" - coded as "size of pills"	Interviewer: Why do you struggle to take your ARV's as prescribed?	"really big" - coded as "pill burden"
"forget a lot" - coded as "difficulty remembering pills"	Participant: Well, the pills are really big. Also, I forget a lot!	
"lot on your mind" - coded as "difficulty remembering pills"	Interviewer: You forget? Tell me a bit more about that?	
"many pills" - coded as "number of pills"	Participant: I guess when you have HIV there is a lot on your mind, so many pills you need to take, you sometimes forget.	"many pills" - coded as "pill burden"
"feeling sick" - coded as "effect of illness"	Interviewer: I see, how often does this happen.	(The rest of the quotes do not fit into any of the predetermined themes).
	Participant: Only on days when I am feeling sick, or on anti-biotics...	

Table 4: using an inductive approach	Read	Code	Interpret
The number of cycles will depend on your personal approach, as well as the amount and depth of the data you have collected. Your cycles may not always follow this scheme exactly, and you may decide to move between phases as you go.	**1st cycle:** Read all your transcripts and field notes. Make some new notes about what you think are the most important themes emerging.	Analyze all the transcripts using appropriate software, or manually. Initially, you can be quite liberal about creating new sub-themes. As you progress, try and use more of your existing themes and create fewer new ones.	Look at the themes you have made. Look at which one's appear most often, and which seem to fit together. You can merge together themes that are very similar, and delete ones that don't seem as important.
	2nd cycle: Read all the sections you have coded, making sure you have put them in the correct categories. Flag the quotes you think are most important.	Analyze the transcripts again, this time restricting yourself to the list of themes you have created (your codebook) - although you can modify it if necessary. You may find that you do not need to recode your later transcripts, if your codebook was well established by this stage.	Again prune your codebook, merging and removing themes as needed. Establish hierarchies of how the themes relate to one another. Think about how the themes you have found can be used to answer your research question. See Box 6.
	3rd cycle: Read the sections you have coded, paying attention to how the quotes can be used together to build your argument.	If you feel that you still need more evidence to build your arguments, you can analyze sections of the data a third time, looking for specific keywords and phrases. At this point, you may wish to quantify certain things, e.g. how many people expressed a certain broad view.	Decide which themes emerged most frequently, and which can be argued most easily. Some or all of these themes will then be the basis of your results and conclusions.

determined that other authors have cited pill burden and a lack of understanding as being an important factor in non-compliance. We can see that there are two examples of these ideas in this extract. Using a deductive method, we would code each of these examples according to the theme that they relate to. In practice, we may have several themes, and each of these themes may in turn be divided into sub-themes. See Box 6 for an example of a way to structure your themes.

The more commonly used inductive approach is different in that you allow themes to emerge from your data. In other words, instead of going and looking for themes that you have found in the literature, you just read your data and try and see what themes people have brought up. An inductive approach generally begins with reading through your transcripts and making notes of all the interesting things that you observe. At the same time you will highlight specific sections of your transcripts, and try and assign them to a specific theme - thus "coding" them. Using this data in conjunction with your field notes will help you identify patterns within your data. The process is illustrated in Table 4.

How do you physically code your transcripts? This can be done either electronically or on paper. If you want to do it electronically, you need to use a form of qualitative research software. There are a few available, and although they are generally rather expensive, your supervisor may have access to such software, and free trials are often available online. This software helps in that you can select portions of your transcript, highlight them and assign them to a specific code. Later, you can call up all the quotes that were assigned to a specific code, and thereby have access to all the data relating to that code. As there are different software packages available, precise details are not included here.

Doing it on paper is more cumbersome, but certainly possible. In this case, you highlight sections of data that you think are important, and can then use different colours for different themes, or label the highlighted sections in a way that you find most convenient. You may want to print several copies of your transcripts and cut out different pieces of text, to allow comparison between all the pieces of text coded in the same way.

Box 6: Sample hierarchy of themes

Pill Burden

- Size of pills
- Number of pills
- Side effects of pills
    - Perceived
    - Actual

Understanding of ARVs

- Importance of compliance
    - Resistance
    - Viral load
- Mechanism of action
    - "Antibiotic" misconception

External social factors

- Difficulty maintaining confidentiality
- Theft of pills
- Difficulty getting to clinic

Write up your study

1. Decide which results to present
2. Describe your results
3. Draw accurate conclusions
4. Choose a method of presentation
5. Finalize and publish

3. Write up your study

1. Decide which results to present

Two issues are important here - firstly, which themes emerged most strongly from your research? After doing a thorough analysis, it should be fairly easy to determine this. Your field notes will help you determine which themes drew your attention early on, and how your final understanding of the research problem has changed from your understanding at the beginning. Critically examining the evolution of your own insights should help you in deciding how your project can help others evolve their insights too. Secondly, look at the frequency of each theme's occurrence, as well as how strong the examples are. It is important to consider both of these factors. You may feel that a specific theme emerged often, but that for some reason it won't be easy to describe. Reasons could be that it was a response that relied upon a leading question, or that it was something that participants never really agreed upon, which could make it difficult for you to present a coherent argument. For this reason it is important to ensure that you pick a theme that was not only common, but that you will be able to describe in a convincing way.

It is also important that you decide how many themes you intend to describe. This could depend on a number of factors, including the complexity of the themes, and the medium in which you intend to present them.Make sure that you will be able to describe them in convincing detail, and give adequate attention to the relevant literature in your review and discussion.

2. Describe your results

The key to success in this section is to use short, simple sentences, and simply report the data you have uncovered without actually discussing its significance. In qualitative research, this can be a difficult task, as even deciding which data to use represents a degree of interpretation. Try as far as possible to let your data speak for itself. This will improve the perceived reliability of your results.

One way of doing this is to make appropriate use of quotes. See Box 7 for an example, and note the clear flow of argument, which proceeds from making a claim to providing clear evidence. On occasion, you may find that the quote does not speak for itself sufficiently, in which case you can either add an additional quote, explain your argument further by referencing your field notes as another source of evidence, or explaining the context of the quote. If you find yourself having to look to other sources, such as the literature, you are probably entering territory that should be reserved for the conclusion/discussion section.

Although this is qualitative research, it may add credibility to your results if you include some numbers, particularly for the most important ideas. Whenever you find yourself writing "many participants felt..." consider whether or not you could actually say exactly *how many*.

Box 7: Extract from sample results section

The concept of pill burden was raised by 13 of the 15 participants. Most often, reference was made to the size of the pills. As one participant indicated:

"I don't know how they expect anybody to swallow such big tablets"

This sentiment was re-iterated several times. An additional factor related to the number of the pills:

"You have to think about your antibiotics, your vitamins, you know!"

However, this specific view point was less unanimously expressed, with more than half of the participants expressing a contradicting view

"It's not the number of tables, I mean I have a pillbox..."

A final factor relating to pill burden was the concept of drug side-effects. This encompassed references to both actual and perceived side-effects ...

Box 8: Extract from sample discussion section

An important finding from our study was that pill burden (specifically, the number and size of pills that patients need to take) was a significant barrier to compliance. This is in accordance with prior research conducted in India (Ramjee et al, 2004), which described pill burden as "a significant battle in the fight against HIV". Our data provides further evidence that this problem needs to be taken seriously, as it represents a barrier across at least two diverse cultural groups. Further research needs to be conducted within more affluent communities to determine the extent to which this phenomenon is present in different socio-economic contexts ...

3. Draw accurate conclusions

This is the most important and most difficult part of writing up research. You will need to ensure that everything you say is meticulously justified. At the same time, if you are too hesitant, you will fail to do your efforts justice. The goal of this section, which is most often referred to as the "discussion" is to summarize your results, and then say what you think they mean. In addition, you need to show how your results tie in with your literature review, and thereby help to position your data within existing knowledge and debate.

The exact way you report your conclusions will depend on the medium in which you intend to present your research (poster, conference presentation or paper). Details on choosing between these methods, and more information on how they can be properly executed, can be found in Document 2. However, whatever method you choose, it is important to have a clear argument in mind. A good place to start is to go back to your research question, and make sure that your argument leads to a conclusion which directly addresses these concerns.

For example, in the case study we used throughout these guidelines, we sought to look at patients' opinions of the most important factors determining ARV compliance. Our conclusion could begin by stating this, and then summarising our data by simply stating which factors emerged most prominently. This will essentially then summarize our results section. We can then state how these findings contrast with the literature, and using these two sources of evidence, draw a broad conclusion. See Box 8.

4. Choose a method of presentation

Depending on what you envisaged for your project, the next step will be to prepare either a manuscript for submission to a journal, or the preparation of a poster or presentation for a conference. See Document 2 for further information.

5. Finalise and publish

Further guidance regarding this process can be found in Document 2.

toolkit

DOCUMENT 4: QUANTITATIVE GUIDELINES

http://www.stockfreeimages.com/

Introduction

At its core, the goal of quantitative research is to quantify the relationship between variables.

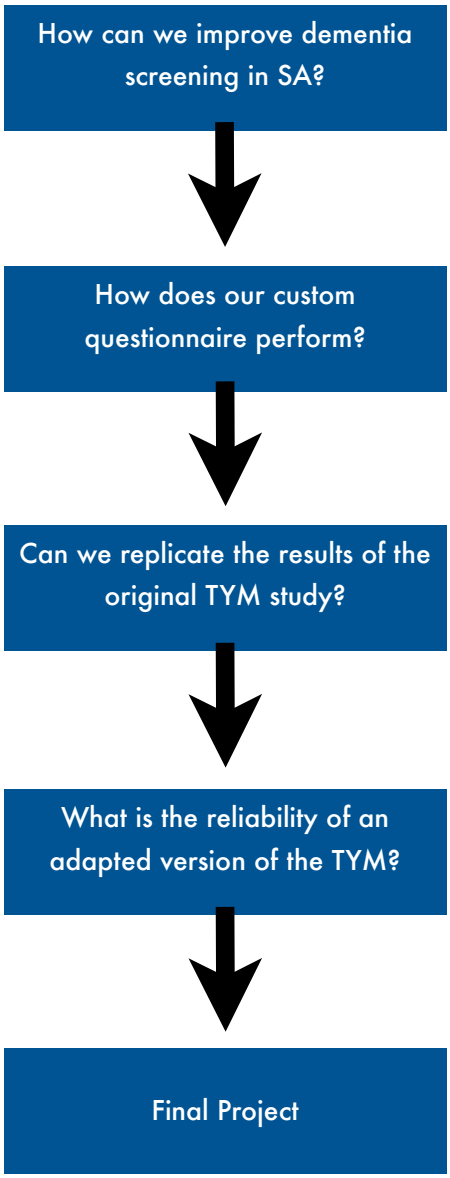
At this point, you would have explored potential avenues of research and concluded that a quantitative approach is most suited to the project you have in mind. Before we get started, it's worth revisiting your research topic and ensuring that you are up for the challenges it presents: Research takes a deceptively large amount of one's time, especially if only done on a part-time basis. The process is made a lot more enjoyable if the topic is truly interesting to the researcher. Early in one's research career, there's a tendency to tackle big problems, or to aim for groundbreaking findings. However, at this stage of your career, should you discover a simple solution to a major problem, or stumble upon a revolutionary therapy, the likelihood that you are mistaken far exceeds that of you managing to hit upon something countless others have missed! As the various 'solutions' to the problem of cold fusion illustrate, doing something no-one else has done is not always a good thing. Be conservative in your goals, and rather aim to complete a neat, well-conducted project with sound methodology. Furthermore, there is no shame in repeating, with minimal modification, a previous study - independent replication is central to the scientific method. Start small - in the future you might find yourself taking on much bigger projects!

contents:

Section summary:

1. The question to be answered
2. Choosing a type of study
3. Decide on a study population
4. Unit of analysis and measuring technique
5. Decide on sample size
6. General considerations

Refining a question:



1. Plan your study

Ask any researcher and they will tell you that the most important part of a project is planning. If everything is set up well, data collection, analysis and interpretation follow smoothly. However, if one has to start analysis on incomplete or incorrectly recorded data, not to mention data that is significantly biased, the process can become laborious and frustrating. In the course of this section you will be guided in the process of planning a simple yet methodologically rigorous study.

1. The question to be answered

Once you have an idea for a research project, you will need to refine it to a hypothesis or question that can feasibly be answered by a study done within your time and budgetary constraints, as well as your limited experience. This does not mean that you have to settle for a simple topic you don't find interesting! Sometimes, a few tweaks can turn a complex question and intricate study into one more fitting for a first time researcher.

The TYM, or *Test Your Memory* project serves as an instructive case study. This study was done by final year medical students, during a five week Family Medicine rotation. Although imperfect, much can be learnt from their experience. The students compared a new dementia screener (the TYM) with the Mini Mental State Examination (MMSE) in order to assess the feasibility of this new tool in the South African context. In the original study, a large group of patients were tested extensively with various cognitive instruments, including the TYM and MMSE.

Put yourselves in the shoes of the students who did that study at the time before they had even thought about the project. Let's say you had heard about, and witnessed, the lack of screening for cognitive impairment among elderly patients in the primary care sector. Due to the time pressure involved with working in a busy day clinic, it may not even cross a physician's mind to screen for early signs of Alzheimer's. So you decide to do a project to address this issue: you are going to design a quick and easy questionnaire that patients can fill out in the waiting room as a screening instrument. However, you soon realize that such a project requires an enormous amount of time, money and expertise. After a brief literature review, and with substantial input from your supervisor, you decide to rather use one of the screening instruments that has already been developed and validated abroad. One particular instrument - the Test Your Memory questionnaire - is appealing since it is self administered, can be scored by a nurse, and has been shown to be sensitive and specific for various types of cognitive impairment in the United Kingdom. You decide to validate it in the South African context. However, you are soon faced with several hurdles - to validate it by repeating the methods used in the original study would be expensive and time consuming. But a few modifications yields a perfectly suitable project: instead of validating it by repeating the original study, you decide to adapt and translate the questionnaire, and then assess its internal reliability and its correlation with the current screening test of choice in South Africa - the Mini Mental State Examination. Whilst you will not necessarily be able to prove that the test performs well

in South Africa, a high reliability and reasonable correlation will warrant further, more detailed study. Furthermore, you can do this in a relatively short period of time and with very little funding.

2. Choose a type of study

Deciding on the type of study you are going to use is vital. Most medical students would agree that one of the topics covered ad nauseam during their training is the study types found in the literature. As such, we won't be going through these in detail. A more comprehensive overview is provided in Box 1 and 2, but the most feasible projects for undergraduate researchers would fall in the category of an observational (rather than interventional) study, be it clinical or epidemiological. Among the clinical studies, due to time constraints and the fact that you need to keep a passing grade, diagnostic studies and case series are the best options. As for the epidemiological group, case-control studies, cross-sectional studies and describing data within a registry can make for good undergraduate projects.

To illustrate why these make for good projects, and how the study type will follow from the research question, consider a student interested in tuberculous meningitis (TBM). Following the advice given above, she decides on an observational design, meaning that the attending physician will manage the patient as he or she sees fit, and that the student will merely be collecting existing data, or performing additional analyses that will not alter the clinical management of the patient. In terms of diagnostic studies, she could assess the sensitivity and specificity of certain clinical features of TBM - say, absent knee and ankle reflexes - by comparing it to the eventual findings on CSF culture results. See Box 3 for more information on gold standards, and why this particular choice won't be a good one. Alternatively, though still a diagnostic test, she could assess the reliability of a clinical finding - for example, the presence of hydrocephalus on CT - by either asking the same physician(s) to read the scans on two occasions (intra-rater reliability) or by having different physicians read the same scans and comparing their interpretations (inter-rater reliability). Should she choose to do a case series, she could collect a number of patients who present with clinical meningitis over the course of a year and describe their clinical presentations or their findings on CSF analysis. In terms of epidemiological designs, she could compare the cell counts and biochemistry of culture positive TBM cases with culture positive 'bacterial' meningitis cases through a retrospective review of laboratory data. As TBM is a notifiable disease, she could also describe the incidence and prevalence among age groups or communities, by using the registry. And there are countless more options, within these relatively 'friendly' designs, that she could pursue

3. Decide on a study population

To a large extent the research question defines the population to be studied. If we think of our TYM case study, it is likely that the population will need to be defined as a specific age group, given that the focus of the test is on detecting cognitive impairment. However, this still leaves a fairly large group from which a study population can be chosen. The details do matter, as this is an area where bias can often enter a project (see Box 4). It will also determine the external validity of the study, i.e. the extent to which results are representative of the 'true' situation in the target population at large. Of course, there is a lot of interaction between the study population and the site of the study, and in order for results to represent a larger community (e.g. the Western Cape) one would likely need several groups (different races, genders, ages, and so on) as well as different sites (Tygerberg, Helderberg, Worcester, and so on). There are numerous ways of identifying your study population, and the different methods have their respective strengths and weaknesses, but it is important that you know how it is going to be done, why it should be done this way, what the biases of the chosen method are and how it affects the validity of your study.

As an example, consider one of the quantitative examples described in Document 2: What is the inter-rater reliability of respiratory findings on auscultation? Here, the study population is the clinicians, as the question we are interested in relates to how their auscultations findings correlate. In addition, we also need to decide on the patients which have to be examined. Regarding the patients, we would prefer to have a good mix of respiratory findings, including normal patients. If we had picked only severe cases, it might result in a higher reliability since the signs may be clearer.

We are not too concerned about generalizing our findings to a larger population of patients with respiratory pathology, therefore it doesn't matter whether we collect them as a set series of consecutive patients seen in a clinic, or from a daily browse through the x-ray database. We are interested, however, in generalizing to physicians, or at least a certain group of clinicians. Describing the agreement

Box 1: Types of studies

Moving from broad groupings to more narrow ones, one can start by dividing studies into either *primary* or *secondary*, with the former representing studies that generate new data and the latter projects that use data already available (meta-analyses, for example). Within *primary research*, three subcategories exist: *basic research*, *clinical research* and *epidemiological research*. Basic research is rarely performed by undergraduate medical students. *Clinical* and *epidemiological* research can be either *experimental* or *observational*. As mentioned before, the *experimental* subtypes of both *clinical* (phases I-IV of clinical trials) and *epidemiological* (interventional studies in the field or on groups) designs are seldom feasible for an undergraduate student. Observational studies, however, provide many opportunities.

Clinical	Epidemiological
Case reports	Cross-sectional studies
Case series	Case-control studies
Diagnostic studies	Cohort studies
Prognostic studies	Description of registry data
Therapy/Drug observational studies	Other (Monitoring, Ecological)

Box 2: Mind your methods

A study's design sets a ceiling of sorts, and no amount of spin or special pleading can salvage a project that has severe methodological flaws. A study assessing the common reasons for rejecting a manuscript submitted for publication, surveyed several editors from top medical journals, as well as a number of Nobel Laureates. Unsurprisingly, study design errors and faulty or poorly described methodology, were by far and away the most common problem. Therefore, spend enough time planning exactly how you will go about doing your research project and get plenty of input from your supervisor, or perhaps even a statistician. Putting in the effort beforehand will make the project much better, but also infinitely less frustrating!

among John, Peter and Bob is rather useless to anyone not working with them! But if they were all registrars in internal medicine, or consultants in pulmonology, we can start inferring the inter-rater reliability of the signs among these subgroups. To comment on the reliability among the broader medical community, we might want to choose a student, a general practitioner, a registrar in internal medicine and a consultant, for example.

At this stage of your research career it may be safer not to try to generalise to a large population, and rather to be more conservative. Remember you need to define your population in terms of person, place and time (who, where and when), and then to select your sample from the same population as that which you have defined in order to avoid selection bias.

4. Unit of analysis and measuring techniques

Deciding on what exactly you will be measuring is of great importance. But it is not always as straightforward as it might seem. To measure the degree to which a CXR is consistent with TB, one would have to decide on which 'signs' to include and how much weight you will give to each. Alternatively, one can ask radiologists to read the CXRs in a more open-ended fashion, but this would complicate your analysis significantly! Think it through carefully: what do you want to measure, what are the options and why are you using this approach?

The method or tool used to make the necessary measurements will also influence your findings, as every test, questionnaire or algorithm comes with its own pros and cons. When in doubt, it's best to stick with the accepted methods, emphasizing the importance of a thorough literature review. Whatever you decide, the procedures should be standardized as much as possible, the ideal being that the same person performs the same test with the same instrument. If more than one person will be used, one should calibrate them with a series of

Box 3: Gold standard

If a new test for tuberculous meningitis is developed, as in the example described under point 2, its sensitivity and specificity will need to be determined. However, in contrast to the theoretical 2x2 tables found in biostatistic textbooks, the amount of 'true' positive and negative cases are not known. Our current best diagnostic test can be used as an approximation of this, meaning that it will serve as the gold standard by which the new test will be judged. The concept of a gold standard is more complex than it might at first seem. Consider the situation where a new test might outperform the current gold standard. How would we establish this? When the new test labels more cases as positive than the gold standard, these could be false positives or cases missed by the gold standard! In practice this issue is often avoided by acknowledging the difference between the current gold standard in the clinical context, and the gold standard used to decide on the value of a new test.

Returning to the example of tuberculous meningitis, one could argue that the gold standard should be the microbiological identification of tuberculous bacilli in the CSF of patients. This is what we normally take as proof of infection in a clinical context, and in fact, many studies use this as a gold standard. But whilst it has a near perfect specificity, it is well known (based on autopsy studies, repeat lumbar punctures and so on) that the yield of microbiological tests for TB in CSF is low. This could potentially result in the above mentioned difficulty with cases that are only positive on the new test. It is because of this that every attempt should be made to maximize the yield of the gold standard. When faced with precisely the issue described above, some researchers have maximized the likelihood of microbiological identification by taking more CSF or doing repeated LPs. In addition, some have used clinical case definitions, classifying patients into various 'degrees of confidence' of a diagnosis of TBM. When choosing a gold standard, make sure it will give you the highest possible sensitivity and specificity, and make every attempt to increase this, rather than simply going with the usual clinical gold standard.

'test runs'. Unless, of course, you are assessing their reliability!

Lastly, a seemingly trivial decision with enormous consequence is the level at which measurements are recorded. Often, one can choose between a metric (the exact value, e.g. an Hb of 13,7), ordinal (one of a set of ranked options, such as normal, high or low) or nominal scale (usually a binary answer to a question, although not all nominal scales are binary, for example, race could be one of a few options that are not ranked and exists as mutually exclusive, all or nothing options). The important point is this: metric values can always be converted to ordinal ones, which in turn can be converted to nominal ones, but the reverse is not true. As such, where possible, record values in 'raw' form, on a metric scale, even though it might need to be converted for analysis.

5. Decide on a sample size

It is important to think about how large your sample needs to be. For the most part it won't be necessary for you to calculate these yourself, if indeed it needs to be calculated for your project at all. But it is an easy way to broadly assess how large a sample you should aim for. It also affords us the opportunity to review some statistics!

Before we get to sample size, we need to cover the concept of 'hypothesis testing'. Although this topic is covered rather extensively in the undergraduate degree, students seem to miss the purpose of the process and hence are unsure about the interpretation of results. Different study designs, measurements and tests will use a slightly different approach, but we'll pick a particular example, as it's the broader concept that you need to understand.

Assume I have developed a new drug for migraine, and I now have to prove that it's more effective than a placebo if I want to have any chance of becoming rich and famous. So I decide to do a trial (and not a very good one): I randomly select 40 students who are known to suffer from migraine. I give 20 of them a pack of placebos and the other 20 get my new drug, called Acephalgia. After a month of use I review their migraine diaries to see how many

Box 4: Guarding against biases

A research project is riddled with potential areas of bias: deciding on the study population, measurement tool(s), data to be captured and statistical test(s), to name but a few. Awareness is the first step towards prevention and expanding your database of biases to be on the lookout for is an important step in your research career. Obviously, detailing all the commonly encountered biases is beyond the scope of this guideline, but a few are worth noting. Perhaps the most common is selection bias, of which numerous types exist. A study may, inadvertently, select the most ill patients, or suffer from a self-selection bias if people are recruited via email. It may also be more overt if inclusion criteria are very stringent, which is what happened with numerous early studies on interventions for ischemic heart disease, which included mostly elderly, white, smoking males with numerous co-morbid conditions. Confounding occurs when the correlation between two variables is actually the result of both variables correlating, independently, with a third variable. The best known example of this is the relationship between alcohol consumption and lung cancer. It turns out both of these correlate well with smoking, so although it would appear that alcohol causes lung cancer, in reality, smoking is causing lung cancer, and it happens that smokers are also more likely to be drinkers! Lastly, both the investigator and the information biases occur when the study is actually being done, the former referring to ways in which investigators may affect results and the latter to erroneous measurement or recording. Assessing a study for potential bias can be a lot of fun!

recurrences they have had after taking medicine for a migraine. The results are as follow:

	Placebo	*Acephalgia*
Recurrence rate	0,6	0,4

Now, there are a few options as to the possible outcomes of my study:

1. My drug works, and my study has shown that it does
2. My drug does nothing, but my study found that it does because of biased recruitment, for example, or
3. My drug does nothing, but because of chance, my study suggests that it does work.

To see whether I designed a bad study leading to bias, or whether there are mistakes in my calculations, and so on, one would look at my methodology, analysis, et cetera. Let's assume the aforementioned are fine, and on that basis we are able to exclude option 2. This leaves us with option 3 - chance. We can test this by calculating the probability that we would get the above mentioned results if there is in fact no difference between the treatments. The latter, of course, is our null hypothesis. There is a chance that, during our testing, we decide that there was a real difference (we reject the null hypothesis) when in fact there wasn't. This is called a Type I error. Similarly, we may decide that the observed effect was the result of chance (we accept the null-hypothesis), when in fact there really was a difference! This is known as a Type II error. And then, of course, we could also accept or reject the null-hypothesis and be correct. These options are summarized in Table 1.

We will touch on some important points regarding the p-value in Box 6, especially the importance of how to interpret it correctly. For the moment, it's clear that we want the chance of making a Type I error to be low, and normally people accept a value of 0.05 or 5%. This predetermined value is labelled as alpha. As for a Type II error, we most certainly do not want to wrongly conclude that my drug doesn't work! We can calculate the chance of us finding an effect if in fact there is one, by calculating what is known as the power of our study. Again, more often than not, you wouldn't need to calculate this yourself, but it will feature in the literature and it might be mentioned by your statistician. There are three things that determine power:

1. The minimum difference we would deem important to detect

	We decide that it's all due to chance	We decide that I made a breakthrough: my drug works
The null hypothesis is true: my drug had no effect	Although sad, this would be the correct decision	Type I error: The chance of us making this mistake is represented by the P value
The null hypothesis is false: my drug did have an effect	Type II error: This is also known as a β-error	Exciting times, and rightfully so: we have correctly decided it works!

Table 1: The possible outcomes of a study

2. The sample size of each group
3. The p-value for which we are aiming.

Generally, a p value of <0.05 is accepted, which practically means that there is only a 5% or lower percent likelihood that our results are due to chance.

Normally, this calculation is done before starting a project, and by setting the power at an acceptable level (often 80%), defining options (1) and (3) and then calculating the sample size needed (2). But we can apply it to my study: the size we detected was a 50% reduction in recurrences, the sample sizes were 20 each and our alpha level is set at 0.05. Plugging this into the formula, we see that the power of my study was around 24%. In other words, at that alpha level, and with those sample sizes, I had a 1/4 chance of finding an effect if there really was one. If I had done the calculation beforehand, and had still gone for a 50% reduction, I would have seen that I needed at least 194 patients to reach a power of 80%.

To avoid being in a similar position, consult with your supervisor regarding the need to do a sample size calculation. It might save you a lot of trouble!

6. General considerations

Once you have decided on your approach, sampling and analysis, you can move forward. The next steps are to set up a timeline for your study, draft your protocol, submit the protocol to your supervisor for comment and feedback, and then apply for ethics approval. This must be done before you can start gathering your data. More information on each of these activities can be found in Document 2 and Document 3 under the relevant headings.

Box 5: Internal validity

Whereas external validity refers to the extent to which a study's results represents the larger population, internal validity is a measure of the degree to which the conclusions are indeed correct. There are several threats to your project's internal validity, but the most important ones are illustrated in the following example: You decide to do a survey on doctors to determine which factors are associated with a greater salary, and you might find that having children is such a factor. Can you conclude that having children increases a doctor's salary? Of course not! Who is to say that people with more money are simply more likely to have children? (Correlation does not equal causation!) Or what if age correlated with both salary and the likelihood of having kids, but you simply failed to evaluate this factor? (Confounding) Even the relatively 'weak' conclusion that there is in fact an association isn't clear: what if you had drawn your sample from doctors entering the day-care facility in Panorama and another group of doctors attending a fertility clinic in Strandfontein? (Selection bias). Logic, and some common sense, will again serve you well here!

Section summary:

1. Introduction

2. Data storage and organisation

3. Assessing a new diagnostic tool

4. Adapting to other designs

Box 6: The high cost of research

As you start your research, you will begin to discover that there are expenses lurking behind every corner. Hopefully, you will have anticipated most of these before starting and have ensured that you have the means to fund your work. But no matter how carefully you plan, there remains a chance that additional expenses will emerge, without which your research cannot be completed. It is vital that you and your supervisor are on the same page regarding this issue. Determine at the outset to what extent you will be able to get support, and agree about how unexpected costs will be covered. Do not be shy to ask around for funding - this is the norm, rather than the exception, in research.

2. Do your study

1. Introduction

The various topics amenable to quantitative analysis, coupled with the myriad of study designs possible to test these, mean that a 'generic' checklist or guide is impossible. However, there are some factors which are common to most studies, and the most important of these will be highlighted below. Thereafter, we will work through how a specific diagnostic study might be done and indicate how things would change if the study design was slightly different.

2. Data storage and organisation

Some simple, but often overlooked, things need to be sorted out before you start with the actual project. Firstly, get organized: label the appropriate files, print your demographics forms, make enough copies of the consent form and ensure you have the stationary you'll need. Set up your database, and decide how and when you will capture the data. Furthermore, you will probably need to design a method of assigning study numbers to your participants, and have a separate database with the demographic details linking each participant to their study number. There is a tendency to want to 'jump in' and get things done, which is understandable. However, a few weeks in, when you are sorting through patients' details written on scraps of papers, or struggling to track down a patient whose consent form got lost, you're going to regret doing so! If ever there was a time to let your perfectionist self run wild, it would be now: take your time.

Should your study involve more open ended responses, such as a person's interpretation of a CT-scan, remember that it doesn't make sense to record them in this fashion in a database. One radiologist might write 'general atrophy' while another might write 'widespread loss of volume', and running statistical tests on 'strings' like this is frustrating. In such a case, you'll have to come up with a way to 'code' the data, and entering the information into a database involves interpreting it to a certain extent.

In terms of the practical aspects of performing your study, the short and simple advice is this: do exactly what you planned to do. If you did a good job in the planning stage, this should be fairly simple. Obviously you have to re-assess periodically and if there are major problems you might need to alter your design. But for the most part, just follow your own instructions! Unless you are working with laboratory samples or doing a retrospective chart review, the process normally involves getting consent from the patient; gathering the required demographic information; performing the index test, or administering the questionnaire, or gathering the data necessary for assessment; gathering additional information, (eg regarding the gold standard), and capturing the data into your database. Once everything has been collected and captured, you (or the statistician) will perform the statistical analyses as planned, and you'll be left with the job of interpreting the information and drawing conclusions!

Box 7: Some important concepts in statistics

Statistics is a field that has always been shrouded in mystery. Seemingly complex mathematics, infinite graphs and a plethora of eponymous tests conspire to create an environment that medical students have tried to avoid by going into a biological science. But if there is one branch of mathematics that one would struggle to escape, for good reason, it is statistics! Thankfully, the level of knowledge required to excel in undergraduate research is more than manageable for most students. This is due, in part, to the fact that you are likely to have help, either from a statistician, or from your supervisor. Meeting with your statistician is a vital step in the planning as well as the analysis parts of your project. This interaction is very much about meeting each other half-way: you need to understand that they lack formal medical training, in most instances, whilst you are lacking in statistical expertise. Know what you are planning to assess, explain this to the statistician, preferably with reference to how others have done it, and do your best to engage with the applicable statistics mentioned by the statistician.

A discussion of all the statistics that might be important is far beyond the scope of these guidelines, but some concepts are almost guaranteed to feature, such as p-values and confidence intervals. P-values have already been mentioned, and are an indication of the *statistical* significance of the result of the statistical analyses done. It should be emphasized that this is *not* an indication of the *clinical* significance. If an antiviral is shown to reduce the duration of illness of the common cold and our tests show that it's *statistically* significant, we have not learnt much. If it reduces it by 3 hours, for instance, it might as well have done nothing, as this is far from being *clinically* worthwhile. Furthermore, a common mistake is to look at the p-value and the value representing the observed effect, and draw conclusions. The confidence interval, on the other hand, is *vitally* important. If *Acephalgia* reduces the risk of recurrence to 0.3 (p=0.01), but the 95% confidence interval stretches from 0.01 to 1.1, we are 95% sure that the 'true' value in the population lies within that range, but it is no more likely to be 0.3 than 1.0, which would indicate no difference!

The specific statistics that apply in your case will depend on various things, most notably the study design used. For example, diagnostic studies are likely to feature sensitivities, specificities and possible predictive values. But statistics is often full of nuances. Consider the assessment of inter-rater reliability: would you use the percentage agreement? The answer, which might not be a surprise, is *no*. This is because a certain amount of agreement is expected purely by chance, so there are specific statistical concepts to account for this, such as *kappa*. If you really think about the calculations involved in your project, and apply some logic and common sense, you would probably arrive at the correct *concept*, even though you might not know the names or formulas involved!

3. Assessing a new diagnostic tool

Let us work through a more specific example. A final year student wishes to assess the performance of a new test for TBM, a 'dipstix' based test that can purportedly diagnose TBM on a drop of CSF. As she followed these guidelines, she spent a lot of time designing her study. She has decided to screen patients presenting with meningitis for inclusion and if they are older than 18, they would be eligible for inclusion. Those with too little CSF to spare some for her project, or those that had received more than three doses of antibiotics would be excluded. Most of the CSF would go to the lab as per normal, and will be used for culture, but she would use some to perform the new test. In consultation with the statistician, she determined exactly how to compare the new test with the culture results.

Starting on her internal medicine rotation, she makes two rounds through the medical emergency unit to look for patients with meningitis. She also uses posters to advertise her project, with her contact details. When she finds an eligible patient, she reviews the chart to check for exclusion criteria and if none are present, discusses the study with the patient. If the patient is willing, the researcher accepts the patient for inclusion and gathers the necessary demographic information on a well structured form. The patient is assigned a study number and the documents are filed. If the LP has not been done yet, she leaves an extra tube in the patient's file with instructions to keep some

Table 2: A sample of reported data

	Methods	Results	Discussion
Patient Selection	All patients over the age of 18 presenting to the medical emergency unit at Tygerberg Academic Hospital with a clinical suspicion of meningitis between March 2011 and November 2011 were screened for inclusion in the study.	A total of 350 patients were screened for inclusion. Of these, 63 were excluded based on the criteria mentioned in Table 1.3. A further 37 were excluded due to inadequate CSF obtained from the laboratory. The demographics of the enrolled 250 patients are given in Table 2.	This study aimed to validate the use of TBStix® in the diagnosis of TBM in the Western Cape, a region with a high prevalence of TB and HIV. The distribution of patient demographics indicate that our sample was representative of the WC at large. Furthermore, we had a low exclusion rate compared to similar studies in Angola and Gabon.
Clinical evaluation	Patients were evaluated by the primary researcher (R.W.) together with a consultant neurologist (J.C.). Patients who met full inclusion criteria without any exclusions (See Table 1.3) were enrolled.	The most common clinical findings were headache (67%), focal neurology (45%) and cognitive impairment (39%). Full clinical features are given in Table 3. Nearly half (44%) of patients had evidence of disseminated TB, with the most common sites (excluding the CNS) being cervical lymph nodes (45%) and the gastrointestinal tract (27%).	As mentioned, a large proportion of patients had evidence of extra pulmonary TB at presentation. This is due, in part, to the late presentation in the community, as well as the fact that TBH is an tertiary referral center. It is worth noting that most of the culture positive cases had presentations warranting empiric anti-tuberculous therapy, calling in to question the use of diagnostic testing in these cases.
Laboratory methods	CSF obtained by the attending physician as part of the workup of the patient was used for the study, provided at least 0.5 ml remained after all the requested tests were done. The newly designed TBStix® were used on the CSF as per the developer's instruction. The results of the diagnostic tests requested by the attending physician were recorded for comparison.	In 56 patients, a positive TB culture was obtained. An alternative diagnosis was confirmed in 110 patients, with cryptococcal meningitis (72%) and viral meningitis (18%) being the most common (See Table 4). The remaining 84 patients did not receive a final diagnosis. TBStix® were positive in 178 patients, of which 53 were culture positive for TB.	Whilst TBStix® was shown to be highly sensitive, the lack of specificity means it can't be used to confirm the diagnosis, and whether or not to start therapy will depend on the clinical features. However, in cases where the clinical suspicion is high, a negative TBStix® result could potentially save the patient from unnecessary treatment and side effects.
Data Management and Analysis	All patients were assigned unique and anonymous study numbers. Patient demographics, the results of the clinical examinations, the laboratory findings as well as the TBStix® results were recorded in Microsoft Excel® and all statistical analyses were done using Stata®.	Using culture as a gold standard, TBStix showed a sensitivity of 94.6% (90.7%-98.5%) and a specificity of 35.6% (24.3%-46.9%). Misdiagnosis frequently occurred with cryptococcal meningitis.	There are several limitations to the current study. Firstly, the gold standard has been called in to question by other authors, who advocate for the use of multifaceted criteria rather than culture results. Secondly, whilst TBStix® has shown promise on pulmonary samples, it could be that CSF is not an appropriate sample on which to run the test.

of the CSF for her. If it has already been done, she contacts the lab to ask for a few drops. She records the result of the dipstix test, as well as all the other information she has gathered, in the appropriate database. Every week, she goes through the laboratory system to check for culture results. Once she has reached her target sample size, she stops, makes sure everything in the database is correct, and contacts the statistician, who analyses the data. Once she gets the report back, she interprets her findings in the light of the literature she reviewed, and draws the correct conclusions!

Table 2 outlines what the methods, results and discussion sections of such research could look like if it was written-up as a paper. This should guide you in thinking about how you will approach reporting on your own research.

Take some time to think about this brief overview of her project. What are some of the problems with her design? Do you think there are any biases? What do you think about her decision to compare the test to culture results? Do you think she could have improved the yield of the cultures? What do you think about her assessment of the culture negative cases? Do you think it is worth investigating these to determine what their final diagnoses were? Do you think she had enough screening sessions? What about the actual LPs – could she have done anything to standardize these?

4. Adapting to other designs

If, rather than assessing a new test, she had decided to evaluate the reliability of laboratory personnel in detecting TB on CSF smears, what would she have done? Much of the design could stay the same, except that there will be no dipstix test. Say, for example, she is interested in the inter-rater reliability, she could have two technicians read a set of slides. They should, obviously, be blinded from one another's interpretation. She could evaluate how much they agree on both positive and negative cases. If, however, she is interested in the intra-rater reliability, she will ask the same technicians to read the slides, and then after a washout period of a few weeks, and assigning the slides new numbers, they will be asked to read them again. With the help of the statistician, she will again analyze the data.

Lastly, if she merely wanted to describe the CSF findings in TBM, only culture positive cases would be included, and the design would change to a case-series. Consent might be waved in this case, but she'd have to see what her supervisor thinks about that!

Section summary:

1. Introduction and literature review

2. Methodology

3. Describe your results

4. Discuss your results

5. Consider international guidelines

6. Choose a method of presentation

7. Finalise and publish

Box 8: A note on style

The ability to write in an appropriate academic style can take a long time to develop, and is often a barrier for younger researchers with little writing experience. One of the most important things to achieve is to write in a simple and clear style, such that the logic of your argument is obvious to the reader. As you gain experience, you may decide to adopt a more complex style, but don't see that as the goal - if you have done good research, it will stand on its own merits. Also take a look at other articles published in the journal you wish to submit to. This can give additional guidance on style.

3. Write-up your study

Regardless of the final product you have in mind – poster, paper, presentation, or all of the above – it's worth writing up your study. As the highest achievement would be a publication in a reputable scientific journal, this is what we will aim for in this section. The same manuscript can, however, be used to prepare for the other options! In general, a manuscript is divided into the following sections: Introduction (or Background), Methods, Results, Discussion and Conclusion.

1. The introduction and literature review

In the introduction, the goal is to discuss why you did your project and why it is important. By providing a focused literature review, you will describe the context of your study, and the results of similar studies. This section generally ends with a description of the question to be answered, and ideally, your hypothesis as to what you will find.

2. Methodology

When discussing the methodology, you need to tell the reader, who knows nothing about your project, enough about how you went about it so that they would be able to replicate the study or the calculations. Imagine that you are describing your project to a clever, critical, but ultimately benign professor. Start by stating the obvious: you got ethical approval. Then run through the questions mentioned in Box 1: What type of study? Who were included? Where did it take place? How many were included? What was measured, how and when? What analyses were done? What instruments were used? Table 2 provides an illustrative example.

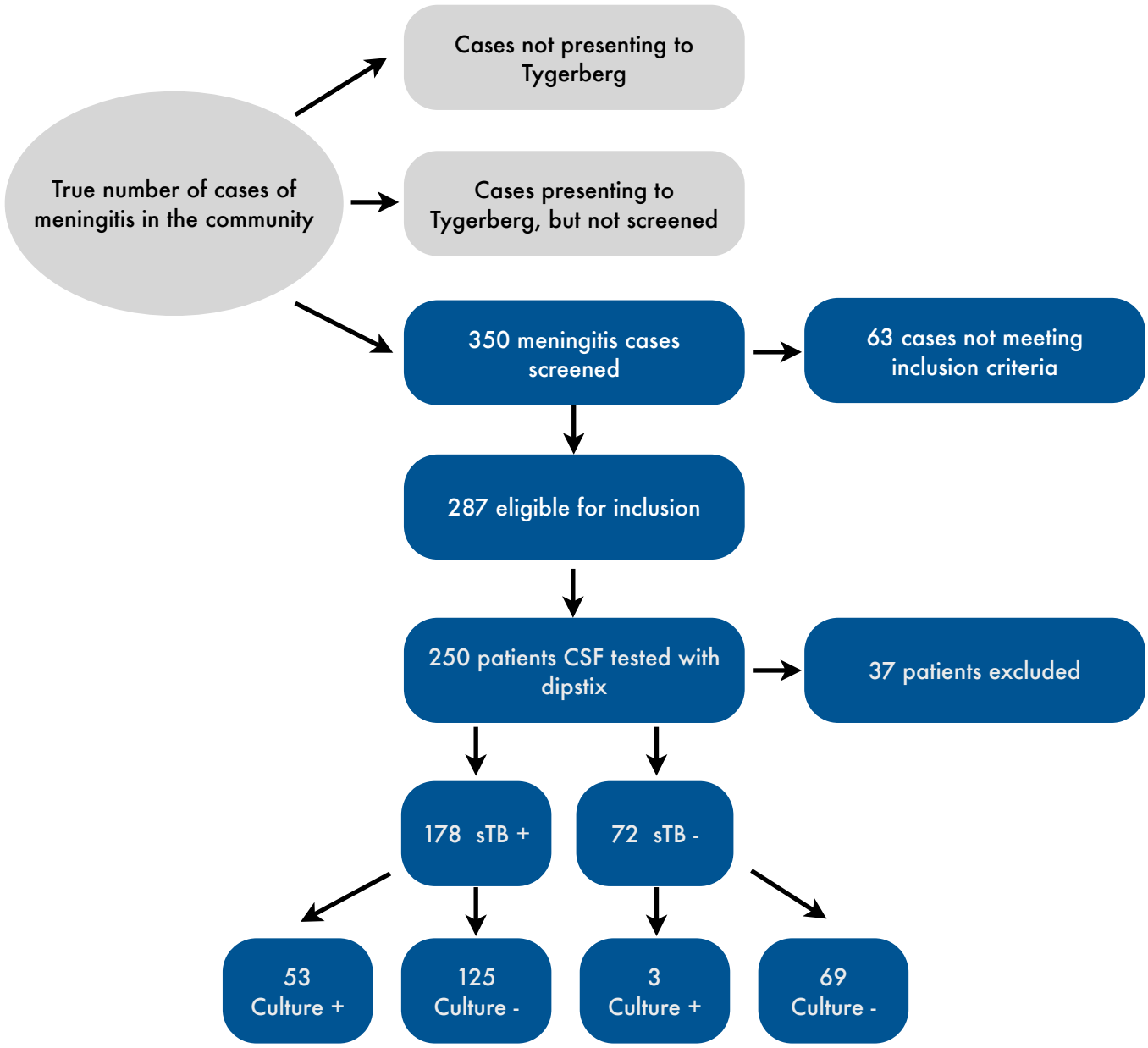
3. Describe your results

In the results section, you will present your findings, but don't interpret them! Just provide the reader with the data. What data? Preferably all of it. This is good scientific practice, and prevents any bias creeping in at this stage. Tables are a lifesaver here, and you don't have to mention everything in both the table and the text, just the important aspects. Irrespective of the statistical analyses that were done, you should get a value, along with a test of significance (p-value), usually accompanied by a confidence interval. Present all of these.

4. Discussion

Finally, you can discuss your findings in the context of the existing literature and draw conclusions. This is the most important and most difficult part of writing up research. You will need to ensure that everything you say is meticulously justified. At the same time, if you are too hesitant, you will fail to do justice to your efforts. The goal of this section is to summarize your results and then say what you think they mean. In addition, you need to show how your results tie in with your literature review, and thereby help to position your data within existing knowledge and debate. Lastly, discuss the limitations of your study and potential avenues for future research.

Box 9: The STARD Diagram



Once you have a working draft, read it as if it is a paper in a journal, i.e. critically appraise it. Ask the questions you'd ask of someone else's findings. Most importantly, ask whether the research question emerges from the literature review, whether the design is appropriate to answer it and if the results support the conclusions.

5. International guidelines

Apart from the specific requirements of the journal that you are considering for publication, most journals require that some study designs be reported in a particular way. The most relevant example of this pertains to studies assessing a diagnostic test, which should tick all the boxes in the STARD-document, which can be found online. The document is easy to understand and to follow. There is one particular aspect that warrants mention here, the 'STARD diagram' (Box 9).

The STARD guideline recommends that one includes a diagram documenting the flow of patients through the study. If we assume that the student above had enough time to recruit a lot of meningitis patients, our diagram might look like the blue-coloured part of box 9. The use of such a diagram should be self evident: it allows the reader to instantly see how many patients were screened, excluded, and how the included patients were classified after the necessary tests. Ideally, one would include the reason for exclusion and if any patients are lost to follow up, these should be added at the appropriate stage. From

the last few blocks, it is easy to derive a 2x2 table to evaluate the sensitivity and specificity.

But it is important to be aware of the aspects that can not be represented in such a diagram, and these are included in grey. Firstly, of all the meningitis patients in the community, only some will present to Tygerberg. Of these, due to imperfect screening, some will slip through. This is why, when moving from our sample to the population at large, we will have to include the interval of values that we are confident the population's value might lie in. Whenever you read a paper reporting on a diagnostic test, run through this mental exercise of adding the 'unseen' elements and consider how these affect the results.

6. Choose a method of presentation

Depending on what you envisaged for your project, the next step will be to prepare either a manuscript for submission to a journal, or the preparation of a poster or presentation for a conference. See Document 2 for further information.

7. Finalise and publish

Once you have completed your research you will want to share it with the scientific community through publication. Further guidance regarding this process can be found in Document 2.

Acknowledgements:

The authors would like to thank the following referees: Prof L van Niekerk and Ms T Oosthuizen.