A profile of Statistics and Research Training of Undergraduate Medical Students at South African Universities

Jean Dommisse

Thesis submitted in accordance with the requirements of the degree

Master in Medical Science

in the Department of Biostatistics, Faculty of Health Sciences,

University of the Free State, Bloemfontein

Supervisor: Prof G Joubert

December 2005

I certify that the thesis hereby submitted by me for the degree M.Med.Sc (Biostatistics) at the University of the Free State is my independent effort and had not previously been submitted for a degree at another university/faculty. I furthermore waive copyright of the thesis in favour of the University of the Free State.

Jean Dommisse

Acknowledgements

To my wife, Christelle, thank you for standing by me in all the hard times, and all the love and support.

Dorothy Cilliers, thank you for the final grammatical touches.

To my supervisor, Gina Joubert, thank you for all the help, and the necessary inspiration that I needed towards the end.

Thank you to all the respondents, for your time and inputs in my studies.

Abstract

Statistics and research methodology are important components of a medical curriculum, since statistical analysis features in the majority of research papers published in medical journals. Medical practitioners need a basic understanding and knowledge of statistics and research principles. Evidence Based Medicine has given an enormous opportunity for statisticians to teach critical appraisal, and to orientate future doctors towards evidence-based practice. Literature on the teaching of statistics and research methodology are available for the United Kingdom, United States and elsewhere in the world but not for South Africa.

It is therefore important to do this study on the profile of research methodology and statistics training for undergraduate medical students at South African universities in terms of the following: (1) What subjects (topics) are medical students taught? (2) Who does the teaching? (3) When is the learning programme / contact sessions taught during the medical students' curriculum? (4) How is the learning programme / contact sessions taught to the students?

I contacted the heads of the eight medical schools in South Africa via email to ask them whether they would give consent for the university to participate in my study. Thereafter I contacted the relevant persons of all the medical schools via email and asked them if they were willing to participate. They needed to complete a questionnaire and checklist. The checklist covered topics taught and the questionnaire the other research questions. The checklist and questionnaire were compiled based on the literature, and tested in a pilot study. One university did not respond, one university does not teach a formal Biostatistics course, one does the Biostatistics course as an elective programme and 5 universities teach the Biostatistics course during the medical curricula. Seven universities completed a checklist and six universities completed the questionnaire. I also requested the learning programme material from the universities to see what the aims and objectives

of their courses are. Five universities supplied me with their learning program materials.

In South Africa the specific statistics or research methodology courses show a vast variety of implementation dates at the different universities. Only one university reinforced the course during the 3rd and 5th year, after it had been taught during the 1st year. For the other, 4 universities teach the course in the 1st year, 1 in the 2nd year and 1 in the 3rd or 4th year, depending on when it is selected as an elective programme. The class sizes vary from 40 to 320 students. Four universities use practical classes and 3 universities use tutors. Three universities use research projects during their medical education. Five of the universities expose the students to Excel, directly in practical classes and indirectly through the research projects that the students must do. The aims and objectives of the South African universities seem on par with what is proposed in the literature.

The persons responsible for the teaching of the statistics / research methodology courses are a doctor (2 universities), statistician (6 universities) and Applied Mathematics lecturer (1 university).

The following topics are taught to the medical students at most universities in South Africa:

(1) Study designs in medical research.
(2) Exploring and presenting data.
(3) Summarising data.
(4) Probability.
(5) Sampling.
(6) Statistical inference.
(7) Analysis of cross tabulation.
(8) Critical reading.

Four universities teach the topic “From sample to population”, “Analysis of the means of small samples”, scatter diagrams and correlations. Only three universities teach the topic of regression. Survival analysis and multiple comparisons are not seen as a core topic in the medical curricula.

Recommendations are made for inclusion of topics in the courses, and for future studies in this field.

Abstrak

Statistiek en navorsingsmetodiek is belangrike komponente van ‘n mediese kurrikulum omdat statistiese analise in die meederheid van navorsingstudies in mediese joernale voorkom. Mediese dokters benodig ‘n basiese begrip en kennis van statistiek en navorsingsbeginsels. Bewysgebaseerde Geneeskunde verskaf ‘n uitstekende onderriggeleentheid in kritiese ontleding vir statistici en om toekomstige dokters te orienteer in die rigting van Bewysgebaseerde praktyk. Literatuur oor die onderrig van statistiek en navorsingsmetodiek is beskikbaar vir die Verenigde Koninkryk, Amerika en elders in die wêreld, maar nie vir Suid-Afrika nie.

Dit is dus belangrik om hierdie studie oor die profiel van navorsingsmetodiek en statistiek onderrig vir voorgraadse mediese studente by Suid-Afrikaanse universiteite in terme van die volgende te doen: (1) Watter onderwerpe word die mediese studente in onderrig? (2) Wie doen die onderrig? (3) Wanneer word die program of kontaksessies aangebied? (4) Hoe word die program of kontaksessies aangebied?

Ek het die hoofde van die agt mediese skole in Suid-Afrika via epos gekontak en hul versoek om hul toestemming te verleen vir die universiteit se deelname aan my studie. Daarna het ek die relevante persone by al die mediese skole gekontak via epos en hulle gevra of hulle bereid sal wees om deel te neem aan my studie. Hulle moes ‘n vraelys en kontrolelys voltooi. Die kontrolelys het die verskeie onderwerpe gedek terwyl die vraelys die ander navorsingsvrae gedek het. Die vraelys en kontrolelys was saamgestel uit die bronne van die literatuur en is deur ‘n loodsstudie getoets. Een universiteit het nie deel geneem aan die studie nie, een universiteit bied nie ‘n formele Biostatistiek kursus nie, een doen dit as deel van ‘n elektiewe program. Die ander 5 universiteite bied almal ‘n Biostatistiek kursus aan gedurende die studente se mediese opleiding. Sewe universiteite het die kontrolelys voltooi en ses die vraelys. Ek het ook die leerprogrammateriaal van die universiteite aangevra om vas te stel wat die doel en doelwitte van hulle kursusse is. Vyf universiteite het dit vir my gestuur.

In Suid-Afrika is daar ‘n groot variasie in datum van implementering van die huidige statistiek / navorsingsmetodiek kursusse by die verskillende universiteite. Slegs een universiteit het aangetoon dat hulle in die derde en vyfde jaar gedeeltes van die kurses herhaal om die studente se kennis te versterk nadat hulle in hul eerste jaar daarmee kennis gemaak het. Vir die ander universiteite word die kursus in die 1ste jaar (4 universiteite), 2de jaar (1 universiteit) en die 3de/4de jaar gedurende die elektiewe program aangebied (1 universiteit). Die klas groottes wissel van 40 tot 320. Vier universiteite bied praktiese klasse aan en 3 maak gebruik van tutors. Drie universiteite verplig studente om navorsingsprojekte te doen. Vyf van die universiteite gee hul mediese studente blootstelling aan Excel, direk in praktiese klasse, of indirek deur navorsingsprojekte. Die doelwitte en oogmerke van die Suid-Afrikaanse universiteite is in lyn met wat voorgestel word in die literatuur.

Die persone wat verantwoordelik is vir die onderrig van die statistiek / navorsingsmetodiek kursusse is ‘n dokter (2 universiteite), statistikus (6 universiteite) en ‘n toegepaste wiskunde dosent (1 universiteit).

Die volgende onderwerpe word by die meeste mediese skole in Suid-Afrika aan die studente aangebied:

(1) Studie-ontwerpe in mediese navorsing.
(2) Verkenning en voorstelling van data.
(3) Opsomming van data.
(4) Waarskynlikheid.
(5) Steekproewe.
(6) Statistiese inferensie.
(7) Analise van kruistabulasie.
(8) Kritiese leeswerk.

Vier universiteite bied die onderwerpe “Steekproef tot populasie”, “Analise van die gemiddeldes van klein steekproewe”, spreidingsdiagramme en korrelasies aan. Slegs drie universiteite bied regressie aan. Oorlewingsanalise en meervoudige vergelykings word nie as deel van die kern van die mediese opleiding gesien nie.

Aanbevelings word gemaak vir die insluiting van sekere onderwerpe in die kursusse, en vir verdere studies in die veld.

Table of Contents

Chapter 1: Introduction and Literature review 1

1.1 Background 1

1.2 From the past to the present (looking at medical statistics, research methodology, evidence -based medicine and continuing professional development) 3

1.2.1 Why should we teach medical statistics and research methodology to undergraduate medical students? 3

1.2.2 What medical statistics and research methodology should we teach undergraduate medical students? 10

1.2.3 Who, when and how should we teach medical statistics and research methodology to undergraduate medical students? 15

1.3 South African regulatory requirements. 21

1.4 Conclusion 23

1.5 Aims and objectives 24

Chapter 2: Research design and methodology 26

2.1 Study design 26

2.2 Sample 26

2.3 Questionnaire design / Measurement 27

2.4 Pilot Study 28

2.5 Data Collection and fieldwork practice 29

2.6 Data management 30

2.7 Analysis of the data 31

2.8 Response Rate 31

2.9 Ethical Aspects 31

Chapter 3: Results 32

3.1 Introduction 32

3.2 Checklist 33

3.2.1	Introduction to medical research	36
3.2.2	Study designs in medical research	36
3.2.3	Exploring and presenting data	37
3.2.4	Summarising data	38
3.2.5	Probability	40
3.2.6	Sampling	41
3.2.7	Clinical measurement	42
3.2.8	From sample to population	43
3.2.9	Statistical inference	44
3.2.10	Analysis of the means of small samples	45
3.2.11	Analysis of cross tabulations	45
3.2.12	Methods based on rank order	46
3.2.13	Multiple comparisons	46
3.2.14	Correlation and regression	47
3.2.15	Other topics	48
3.2.16	Additional topics	48
3.3	Questionnaire	49
3.4	Learning Programme Material	56
3.5	Summary	57

Chapter 4: Discussion 60

4.1	Introduction	60
4.2	Checklist	60
4.2.1	Study designs in medical research	60
4.2.2	Exploring and presenting data	62
4.2.3	Summarising data	63
4.2.4	Probability	65
4.2.5	Sampling	66
4.2.6	Clinical measurement	67
4.2.7	From sample to population	68
4.2.8	Statistical inference	69
4.2.9	Analysis of the means of small samples	71

4.2.10	Analysis of cross tabulations	72
4.2.11	Methods based on rank order	73
4.2.12	Multiple comparisons	73
4.2.13	Correlation and regression	74
4.2.14	Other topics	75
4.2.15	Additional topics	76
4.3	Questionnaire	78
4.4	Learning Programme Material	82
4.5	Shortcomings of the study	83

Chapter 5: Conclusion & Recommendations **85**

5.1	Conclusion	85
5.2	Recommendations	86

References **88**

Appendices

A: Checklist	93
B: Questionnaire	103
C: Consent form (Afrikaans) Heads of Schools	107
D: Consent form (English) Heads of Schools	108
E: Consent form respondents	109

List of Tables

Table 1.1	Aims and objectives of General Medical Council towards the teaching of statistics - 1993	6
Table 1.2	Outcomes and objectives of medical students after successful completion of medical curriculum	9
Table 1.3	Important topics for medical doctors to know in 1978	11
Table 1.4	Topics that were typically covered in the required Biostatistics course at medical schools in the United States	13
Table 1.5	Characteristics of students with different learning approaches	20
Table 3.1	Summary of who teaches what topics	34
Table 3.2	Summary of when the topics are taught	35
Table 3.3	Presentation of topics regarding study designs in medical research	36
Table 3.4	Presentation of topics regarding exploring and presenting data	37
Table 3.5	Presentation of topics regarding summarising data	38
Table 3.6	Presentation of topics regarding probability	40
Table 3.7	Presentation of topics regarding sampling	41
Table 3.8	Presentation of topics regarding Clinical Measurement	42
Table 3.9	Presentation of topics regarding from sample to population	43
Table 3.10	Presentation of topics regarding statistical inference	44
Table 3.11	Presentation of topics regarding analysis of the means of small samples	45
Table 3.12	Presentation of topics regarding the analysis of cross tabulations	45
Table 3.13	Presentation of topics regarding the methods based on rank order	46
Table 3.14	Presentation of topics regarding multiple comparisons	46
Table 3.15	Presentation of topics regarding correlation and regression	47
Table 3.16	Presentation of topics regarding other topics	48
Table 3.17	Number of universities that does groupwork, practical classes, uses tutors and does a research project	50

Table 3.18	Problems encountered in the statistical / research course	52
Table 3.19	Computer programs students are exposed to	53
Table 3.20	Respondents' opinions regarding student views on the research methodology / statistics learning programme	53
Table 3.21	Are the statistics and research methodology taught as separate modules or an integrated module	54
Table 3.22	Stages of implementation of statistics in other modules	54
Table 3.23	Suggestions of improvements	55

Chapter 1

Introduction and literature review

1.1 Background

Florence Nightingale and John Snow applied statistical methods in medical research more than 150 years ago. Florence Nightingale improved the methods to construct mortality tables. She was a fellow of the Royal Statistical Society and an honorary member of the American Statistical Association. John Snow applied simple statistical methods, about the same time; to support his theory that contaminated water was the source of a London cholera epidemic in 1854. Statistics is now an integral part of most medical research projects (Sprent, 2003:522).

Statistics is a very important component of a medical curriculum, since statistical analysis features in the majority of papers published in medical journals. Most medical practioners need a basic understanding of statistical principles, not to mention statistical techniques (Sprent, 2003:522).

Any teacher involved in the teaching of medical undergraduates will know how difficult it is to persuade students that an understanding of medical statistics is necessary, because their principal focus is on the clinical skills they need when faced with their first patients. There are two reasons why we educate medical undergraduates in medical statistics:

1. They will need to interpret research results, as well as understand the implications thereof for clinical practice, throughout their careers, before and after they qualify.
2. They may need to conduct statistical analysis themselves, for instance when analysing the results of a project, which is part of a preclinical degree or a Master's Degree Thesis.

The increased focus on Evidence Based Medicine (EBM) denotes that the first reason has become much more important than the second (Sterne, 2002:988).

Not many medical practitioners conduct medical research, but if they pride themselves on being up to date, they will definitely be consumers of medical research. Continuing Professional Development (CPD) in South Africa has also made it compulsory for doctors to keep up with the latest medical research. This is encouraging because it is the responsibility of medical practioners to discern good research studies from bad, to be able to verify whether the conclusions of a study are valid and to understand the limitations of such studies (Campbell, 2002:1).

Since 1994, all eight medical schools in South Africa have developed and introduced slightly different curricula. Two of them offer a 5-year undergraduate medical programme, one offers a 5½-year programme, and the rest are still offering 6-year programmes. Some follow a fully Problem Based Learning (PBL) approach, some use a hybrid, integrated model and some follow the old classical pre-clinical and clinical type programmes (McKimm and Jollie, 2004:2).

Our aim as educators should be to teach what can be termed as 'statistical thinking' rather than attempting to transform medical students into statisticians. Another important aspect that comes to the forefront in the 21st century is that of not **what** we teach the medical students but rather **how** we teach them. It is important that 21st century doctors are equipped with sufficient critical appraisal skills to assess online journal articles and independent web sites of more doubtful origin and reliability (Palmer, 2002:995-996).

1.2 From the past to the present (looking at medical statistics, research methodology, evidence-based medicine and continuing professional development)

1.2.1 Why should we teach medical statistics and research methodology to undergraduate medical students?

Programs of statistical education for non-statistics majors at universities is intended to serve one of the following three purposes:

(1) Statistical literacy education for the future citizens who are to become the "consumers" of statistics, expected to read statistical data intelligently and think statistically in the information society.

(2) Training elementary and secondary schools' "teachers of statistics".

(3) Teaching statistics and statistical methods for the future "users" of statistical methods in their fields of application: sciences, technology, industry, medicine, business, government, and other (Ito, 2001: 1).

Medical students have a vast amount of information resources at their fingertips, yet there is great uncertainty about how to find the right article to read, and even more uncertainty about how to interpret the data in a paper (McLucas, 2003: 1).

In 1948 the British Medical Association (BMA) Curriculum Committee recommended the introduction of statistics into the medical curriculum, but little was done about this recommendation for at least 20 years (Morris, 2002:970).

Since 1967, when the General Medical Council's *Recommendations as to Basic Medical Education* stated that the undergraduate medical curriculum should include instruction in statistics and biometric methods, it has been generally accepted that this topic should be part of medical training (Wakeford, 1980:73).

In a 1970's General Medical Council report it was found that 29 out of 34 schools in the United Kingdom were already teaching medical statistics, mainly in the pre-clinical stage of the curriculum (Morris, 2002:971).

In 1977 an article was published by Hunponu-Wusu, that stressed the importance of medical statistics in the training of all health workers and the early involvement with the subject was advocated during the training of these health personnel. Four important areas of medical statistics were identified in 1977:

1. The study of specific rates.
2. Evaluation of clinical drug trials.
3. The assessment of factors affecting health or disease.
4. The establishment of new avenues of medical research (Hunponu-Wusu, 1977:351).

A conference that addressed teaching and learning statistics was held at the University of Manchester Medical School in 1979, in conjunction with the Royal Society of Medicine. The 1979 *Draft Recommendations* re-emphasized the importance of statistical methods. Much argument still occurred as to whether the teaching of statistics should be arranged as a separate subject in its own right, or in conjunction with other subjects. According to Wakeford, the needs of medical students are much broader than just the analytic tools of statistics, but the requirements should focus more on the exact measurement and the communication of results (Wakeford, 1980:73).

When the annual meetings of the United Kingdom medical statistics teachers were introduced in 1980, regular updated compilations of the aims, objectives and content of teaching courses in each school were submitted by David Appleton. The most recent compilation of the "blue book" took place in 1996, but some information was up to three years old. At this time, initiatives to promote Evidence Based Medicine (EBM) were about to be realised (Morris, 2002:971).

By 1987, the General Medical Council (GMC) Education Committee was recommending how, rather than whether, the subject should be taught, and integration with other material taught in the curriculum was emphasized (Morris, 2002:970).

The reasons given for teaching statistics to medical students in the 1990s were that:

- Students should be able to perform a critical assessment of medical literature.
- It is unethical to carry out poorly designed studies on human subjects.
- A doctor might draw disastrous conclusions from a clinical experience because he has no concept of scientific method, and believes that the handling of evidence and statistics needs no expertise (Appleton, 1990:1013).

The reason why statistics, and thus medical statistics, is important, as Appleton shows, is the alarming ease with which one comes across poor papers. This raises an ethical issue: Those who perform and report such research should surely be better educated in research methodology, as should the adjudicator who allows the papers to be published (Appleton, 1990:1014).

The above argument for teaching medical statistics, provided by Appleton, therefore contains real substance, but does the solution lie in the teaching of medical statistics (Appleton, 1990:1014)?

The role of statistics in medical research starts at the planning stage of a clinical trial or laboratory experiment to establish the design and size of an experiment that will ensure a good prospect of detecting effects of clinical or scientific interest (Sprent, 2003: 523). Working with statistics involves using statistical methods to summarise data and using statistical procedures to reach certain conclusions that can be applied to patient care and public health planning (Dawson-Saunders and Trapp, 1994:2). Journals for doctors are full of statistical material of this sort, as well as the findings of individual research studies. Statistical issues are implicit in all clinical practice when making diagnoses and choosing an appropriate treatment (Altman, 1991:3).

In 1993 the General Medical Council (GMC) published *"Tomorrow's Doctors"* which proved a catalyst for changes in the medical curriculum. Table 1.1 shows the suggested aims and objectives of teaching statistics to medical students.

Table 1.1: Aims and objectives of General Medical Council towards the teaching of statistics (General Medical Council, 1993)

AIMS:

- To produce doctors with increased skills in diagnosis, prognosis and treatment.
- To educate doctors to be competent to interpret data presented in the press, pharmaceutical literature and professional journals.

OBJECTIVES (students should be able to):

- Understand variability and how to assess it.
- Appreciate the value of medical statistics and the limitations of their knowledge, hence when to request professional statistical advice.
- Understand methods of estimating and the meaning of confidence intervals.
- Understand the methods used when making comparisons, and how these results are presented.
- Reason sensibly about problems involving numbers.
- Assess critically the sources and validity of data.

Recent informal reviews tend to show that medical statistics training has also undergone changes following the publication of *Tomorrow's Doctors*. These tend to include:

- Much more emphasis on critical appraisal.
- More emphasis on concepts and less on techniques.

- Less emphasis on formal lectures.
- More variety in teaching e.g. small groups, workshops etc. (General Medical Council, 2003).

Medical practice is changing, and the change, which involves the more effective use of medical literature as a guidance tool, has even been called a paradigm shift (Evidence-Based Medicine Working Group, 1992:2420-2424).

The evidence-based medicine (EBM) movement assumed its label in the early 1990's when a series of critical appraisal guides were published in the *Journal of the American Medical Association* (Morris, 2002:969).

Evidence-based medicine has been defined as "the process of systematically finding, appraising, and using contemporaneous research findings as the basis for clinical decisions" (Straus and McAlister, 2000:837).

Evidence-based medicine aims to make medical decision-making more deliberate and methodical. Most descriptions of the evidence-based approach contain the following four steps:

1. Formulating a question for research.
2. Searching the medical literature.
3. Critical appraisal of the medical literature.
4. Integrating research into clinical practice (McLucas, 2003: 1).

EBM has given an enormous opportunity for statisticians to teach critical appraisal, and to orientate future doctors towards evidence-based practice. This advantage should outweigh the disadvantage of loss of our discipline's identity in modern curricula (Morris, 2002:969).

The practice of evidence-based medicine requires skills that have not been taught in traditional medical undergraduate programmes. To effectively teach these skills to medical students, revisions to the medical curriculum and faculty organisation need to be made. Adequate educational resources to facilitate learning and the use of information technology are also prerequisites. As the medical community has increasingly recognised the importance of EBM, the profession itself must also address some of these same issues in its continuing medical education programmes (Hazlett, 1998:183).

Difficulties that have been encountered in teaching evidence-based medicine include the following:

1. People like quick and easy answers.
2. For many clinical questions, high quality evidence is lacking.
3. The concepts of EBM are met with scepticism by many faculty members who are therefore unenthusiastic about modifying their teaching and practice in accordance with its dictates (Evidence-Based Medicine Working Group, 1992:2420-2424).

Table 1.2 shows the suggested outcomes and objectives medical students after the successful completion of the medical curriculum according to the 2003 General Medical Council (GMC) publication of "*Tomorrow's Doctors*". Note the mentioning of the statistical knowledge and skills that need to be gained.

Table 1.2: Outcomes and objectives of medical students after the successful completion of medical curriculum (General Medical Council, 2003)

OUTCOMES:

- Be able to gain, assess, apply and integrate new knowledge and have the ability to adapt to changing circumstances throughout the medical student's professional life.
- Be willing to take part in continuing professional development (CPD) to ensure that they maintain high levels of clinical competence and knowledge.

OBJECTIVES:

- Use research skills to develop greater understanding and to influence their practice.
- Solve problems.
- Analyse and use numerical data.
- Take account of medical ethics when making decisions.
- Manage their learning needs.
- Prioritise tasks effectively.
- Reflect on practice, be self-critical and carry out an audit of their own work and that of others.
- Follow the principles of risk management when they practice.

Will it help the medical students or the medical environment if we cram the students' heads full of statistical theories and formulae, but we do not demonstrate to them how to use statistics to make important decisions?

Students do need to be able to critically appraise medical research, and will need to continue to be able to do so throughout their clinical careers, but it is not necessary that they are able to analyse data themselves. If the need arises to do so, when research is undertaken, then the basic knowledge acquired on the principles of statistical methods can be developed and built on as required (Astin *et al*., 2002:1005).

1.2.2 What medical statistics and research methodology should we teach undergraduate medical students?

A reader of medical literature who is acquainted with some simple descriptive statistics (percentages, means and standard deviation) has full statistical access to 58% of the articles in the New England Journal of Medicine. Understanding t-tests increases this access to 67%. The addition of contingency tables gives complete statistical access to 73% of the articles. Familiarity with each additional statistical method gradually increases the percentage of accessible articles (Cheatham, 2000:585).

The most important reason to teach specific topics to medical students must be the relevance of these topics to be included in the curriculum (Simpson, 1995: 202).

In the next few paragraphs important topics are listed. The first list is a list of what was important nearly 30 years ago, and was included to see if changes took place. The second list by Dixon was compiled out of research done on what topics were covered by 20 courses taught at medical schools in the United Kingdom during 1990/91. The third list by Cheatham was compiled out of research done on what are deemed as important topics in 62 courses at medical schools in the United States of America during 1987/88. The fourth list by Looney *et al.* was compiled out of research done on what topics were typically covered in required Biostatistics courses at 125 medical schools in the United States of America during 1992/93. The fifth is a list of topics that was covered by a research methodology workshop in Iran during 2001. Because literature was limited only these are listed.

Table 1.3 lists what was regarded as important knowledge for medical doctors to possess in 1978.

Table 1.3: Important topics for medical doctors to know in 1978 (Wakeford, 1980:73)

- Random and non-random variation.
- Relationship of a sample to a population.
- The nature of a sample limits the inferences that could be validly drawn.
- Few practical statistical methods and tests:
    - standard deviation
    - standard error of the mean
    - probability limits
    - confidence limits,
    - chi-squared test,
    - probability of difference between two percentages.

The following is a twelve-topic course recommended by Dixon (1994:64), a lecturer at the Department of Public Health Medicine at the University of Sheffield Medical School in the United Kingdom.

1.) Data Description: Rates and proportions, types of data, frequency distributions, histograms, median, quantiles and/or percentiles, mean, variance, standard deviation.

2.) Clinical measurements: Repeatability and precision in measurements, comparing two methods of measurement, sensitivity and specificity, survival data.

3.) Probability and decision-making: Properties of probability, Bayes' theorem.

4.) From sample to population: Properties of the normal distribution, reference ranges, standard error of a sample mean, confidence intervals, standard error of a proportion, standard error of the difference between means, standard error of the difference between proportions, sample size for an estimate, Binomial distribution, Poisson distribution.

5.) Statistical inference: Testing a hypothesis, the null hypothesis, the p-value, significance levels, one- and two-sided tests of significance, statistical power, degrees of freedom, confidence intervals rather than p-values, power/sample size, exploratory data analysis.

6.) Design: Cross-sectional surveys, cohort studies, case-control studies, randomised controlled clinical trial, interim analysis and/or sequential trials, association and causality.

7.) Analysis of the means of small samples: t distribution, one sample t method, means of two independent samples.

8.) The analysis of cross-tabulations: chi-squared tests, Fisher's exact test, Yates' continuity correction for a 2 by 2 table, McNemar's test for matched samples.

9.) Methods based on rank order: Mann-Whitney U test, Wilcoxon matched pairs test, Spearman's rank correlation coefficient, and Kendall's rank correlation coefficient.

10.) Vital Statistics: Rates, standardisation of rates.

11.) Correlation and regression: Scatter diagrams, correlation, confidence interval, method of least squares, regression, multiple regression.

12.) Other topics: Choosing the statistical method, critical reading.

The following is a 12-month course recommended by Cheatham (2000:586), from the Department of Surgical Education, Orlando Regional Medical Center in Orlando, Florida.

- **Month 1** – Basic Statistical Theory: Hypothesis Testing, type I/II errors, significance levels, power, one-tailed vs. two-tailed testing.
- **Month 2** – Types of Data: Discrete, continuous, independent, dependent, normal vs. non-normal distributions.
- **Month 3** – Common Descriptive Statistics: Mean, median, mode, standard deviation, standard error, variance, range, bias, precision, confidence intervals, proportions, percentages, odds ratios.
- **Month 4** – Discrete Data Analysis: Contingency tables, Chi-square test, Yates correction, Fisher's exact test.

- **Month 5 –** Continuous Data Analysis: Student's t test, Wilcoxon signed-ranks test, and Mann-Whitney U test.
- **Month 6 –** Multiple Comparisons: ANOVA, Bonferroni adjustment, meta-analysis, post hoc comparisons.
- **Month 7 –** Correlation: Pearson's product-moment correlation.
- **Month 8 –** Regression: Simple regression, multiple regression, logistic regression.
- **Month 9 –** Epidemiologic Tests: Sensitivity, specificity, positive predictive value, negative predictive value, accuracy, prevalence, incidence, mortality rates, probability, receiver operating curves.
- **Month 10 –** Survival Analysis: Kaplan-Meier curves, life-table analysis, and actuarial analysis.
- **Month 11 –** Study Design: Observational vs. experimental, retrospective vs. prospective, longitudinal vs. prevalence studies, controlled vs. uncontrolled, sample size, power analysis, investigational review boards.
- **Month 12 –** Study Quality: Statistical bias, "intent to treat", inappropriate statistical testing.

Table 1.4 contains a list of topics that were typically covered in the required Biostatistics courses at medical schools in the United States during the 1992, 1993 period according to a study done by Looney *et al.* (1998: 93).

Table 1.4: Topics that were typically covered in the required Biostatistics courses at medical schools in the United States (Looney *et al.*, 1998:93)

2. p values
3. Hypothesis testing
4. Interpretation of confidence limits
5. Descriptive statistics
6. F-test
7. Case-control studies
8. Incidence and prevalence
9. Cohort studies

Table 1.4-continue: Topics that were typically covered in the required Biostatistics courses at medical schools in the United States (Looney *et al.*, 1998:93)

10. Frequency distributions
11. Rates
12. Normal distribution
13. Central tendency
14. Randomised clinical trials
15. Study design characteristics
16. Chi-squared test
17. Correlation
18. Variability
19. Construction of confidence limits
20. Descriptive studies
21. Probability
22. Cross-sectional studies
23. Characteristics of diagnostic tests
24. Interpretation of diagnostic tests
25. Interpretation of tables and graphs
26. Scales of measurement
27. Linear regression
28. Measurement issues
29. Adjusted rates
30. Power analysis
31. Binomial distribution
32. ANOVA
33. Construction of tables and graphs
34. Multiple comparisons
35. Wilcoxon Mann-Whitney

According to a medical education research methodology workshop during 2001 at Shaheed Beheshti University of Medical Sciences and Health Services in Iran the following twelve learning modules were marked as important:

1. Selection of a research topic related to a medical and health-related system.
2. Review of literature.
3. Statement of the problem.
4. Statement of research hypothesis/question(s).
5. Identification of variables under study.
6. Selection/identification of measurement instruments.
7. Definition of population and sample under study.
8. Selection of appropriate research method.
9. Selection of data collection procedure.
10. Data processing and statistical tabulation.
11. Selecting appropriate statistical analysis.
12. Interpretation of results and preparing report (Bazargan, 2002:2).

1.2.3 Who, when and how should we teach medical statistics and research methodology to undergraduate medical students?

The way in which we teach medical statistics needs to match as closely as possible the way in which our students will deal with the results of statistical analysis (Sterne, 2002:989). The statistical training of medical researchers usually begins with an introductory undergraduate statistics course, which is possibly followed by a course or two, while they attend medical school (Stangl, 2002: 1).

According to Wakeford (1980:73-54) the following is a summary of what was seen as important regarding whom, when and how medical students were taught in 1974/5. The position in the United Kingdom and Eire during 1974/5 was that not all the schools taught statistics. Of the thirty schools, thirteen had run separate courses on 'medical statistics' or 'statistics' as the sole means of teaching statistics. Six had supplemented such courses with further statistics lectures, provided by other disciplines. The vast majority of the teaching then occurred before the clinical part of the course. The amount of teaching devoted to statistics varied enormously, ranging from around 3 to over 50 hours. The classic statistics course was between 20 and 30 hours and occurred during the first year of their course. In 18 schools the teaching was under the control of a statistician, or biostatistician. Twelve schools used staff members from other disciplines to teach statistics.

At the London Medical Schools, statistics had been a compulsory subject for several years. Each medical school had its own curriculum for statistics. In 8 schools the teaching was under the management of a professional statistician. The classic course consisted of about 21 hours of statistical teaching (Wakeford, 1980:73-54).

The situation in North America was that 72% of schools 'required' Biostatistics training and an additional 14% indicated that such training was available. The classic course involved 21 hours of statistical teaching (Wakeford, 1980:73-75).

According to Wakeford (1980:73-54) the following is a summary of what was regarded as important subject content of when, how and by whom statistics should be taught in 1978. If statistics was taught early in the course, students could not see the relevance of the subject. The other side of the coin is, however, if statistics formed part of the course during the later years, students found 'clinical medicine' more attractive, and were therefore not interested in statistics. The answer, thus, lies in the integration of statistics with other courses and not a separate course dedicated to statistics.

It is evident that a standard first year course of between 10 and 20 lectures, can be counter-productive and will not provide the medical students with a lasting awareness of medical statistics. The ideas that one needs to convey to the students in an introductory course must focus on the more fundamental issues. The actual mechanics of specific tests are trivial at an early stage, but can easily be supplemented at a later stage, when the students are more motivated, because of their own experiences of genuine clinical problems, where they had to apply statistical methods (Murray, 1990:1063-1064).

In the study of Looney *et al.* (1998: 92-93) on Biostatistics requirements in United States medical schools, it was found that in those courses that required Biostatistics (74 schools out of 100 schools that completed the information and there were 125 schools in total), the numbers of hours of required instruction in Biostatistics varied widely across medical schools, with a range from two to 48 hours (median = 20; 69 schools responding). At approximately two thirds of the schools requiring Biostatistics (65%), the respondents felt that the instructional time allocated to the course was sufficient to cover the course material.

When teaching should take place, is a difficult question, but a method that succeeds is an early pre-clinical course (so that the 'jargon' of statistics is familiar) and a reinforced course during the latter studies (especially in courses of epidemiology and health statistics) and the clinical years (Clayden, 1990:1033). The norm is, often, to teach the statistics course during the first semester, when the medical student still exists in a "vacuum of medical knowledge". The lecturers should thus be able to provide the student with a frame of reference and should hence possess knowledge of, and insight into, medical literature (Stander, 1999).

The skills, to appraise and interpret medical research, are needed throughout the undergraduate course, particularly with regards to clinical medicine. It therefore makes sense to master the relevant statistical concepts as early as possible. It is important, however, that these skills are used and reinforced throughout the course, otherwise they may just be learned for the first year exams, and then forgotten. It is therefore highly desirable for critical appraisal skills and statistical interpretation to be integrated into the rest of the undergraduate curriculum (Astin *et al*., 2002:1005).

In the study of Looney *et al.* (1998: 92-93) on Biostatistics requirements in U.S. medical schools, it was found that in those courses that required Biostatistics (74 schools), 41 schools (55%) offered it during the first year.

Twenty-four schools (32%) offered it during the second year, and 4 schools (5%) began the course in the first year and continued it until the second. For the other five schools one offered it in the third year and four in the fourth year.

Lecturers from a number of different disciplines teach medical statistics. **Who** the teachers of statistics should be depends on **what** is taught. With this in mind, medical statistics can be divided into three broad types (Clayden, 1990:1032 - 1033):

- A course, which consists entirely of statistical theory.
    - Theoretical statisticians with a strong theoretical interest invariably deliver such courses. This type of course is likely to be unattractive to the large medical undergraduate audiences most of us are faced with.

- Medical statistics courses, which involve only the applications of formulae and facts.
    - Non-statistically trained teachers may well be in charge of this second, more practically oriented type of course, which may be popular with students, as they put medical statistics in its proper context. Colleagues in the medical school may view such a course as an integrating, or more likely, a supporting feature of the curriculum.

- Medical statistics course, which include both theory and practice.
    - This intermediate type of medical statistics course is one that involves a theoretical component, often followed by an opportunity to practice what has been recently, preferably immediately, been taught. This type of course is likely to involve both medically and non-medically qualified (including statistically qualified) teachers.

The lecturing of service courses are, unfortunately, often left in the hands of junior lecturers, while the experienced lecturers teach the more challenging mathematical courses. This often results in a disinterested teacher, who also exists in a “medical vacuum”, and who teaches students who also function in a “medical vacuum”. Students consequently merely try to pass a course for which they see no purpose in the medical profession (Stander, 1999).

In the study of Looney *et al.* (1998: 92-93) on Biostatistics requirements in U.S. medical schools, it was found that in those courses that required Biostatistics (74 schools), at 41 schools (55%) a PhD faculty member had primary responsibility for the course, at 7 schools (9%) a MD faculty member was in charge, at 17 school (23%) a MD and a PhD co-directed the course and at 9 schools (12%) a masters-level faculty member directed or co-directed the course.

Statistics involves the learning of new skills, almost a new language, and thus a more interactive form of teaching is necessary, an approach where problems and methods can be discussed, and feedback can be given to students to help ensure that their understanding is correct. Small group teaching sessions are therefore an appropriate and necessary format, although some difficulties are experienced with this approach, due to the limited supply of qualified staff, and suitable tutor rooms to conduct 15 or more tutorial sessions (Astin *et al.*, 2002:1005).

In the study of Looney *et al.* (1998: 92-93) on Biostatistics requirements in U.S. medical schools, it was found that in those Biostatistics courses that were required, the numbers of students varied greatly, from a minimum of 40 to a maximum of 265 (median = 134; 72 schools responding). Almost all (89%) of the medical schools that required an introductory Biostatistics course still used lecturers as a method of instruction in that course; lecturers were the only method used at 22% of these schools. Various other methods of instruction were also used, including small-group exercises (55%), take-home exercises (28%), computer tutorials (27%), and students’ presentations during class time (16%). Only 19% of the required courses provided instructions on the use of computers. Of these 19%, 50% covered statistical software, 43% covered database management software, 29% covered word processing software, 29% covered spreadsheets, and 21% covered computerized library searches.

The approaches that students take to study need to be considered when designing medical curricula, so that optimal patterns of learning behaviour are rewarded. University teachers hope that students will adopt a "deep-learning" approach (Table 1.5) with the aim of gaining understanding by reading widely, by asking questions and by exploring new concepts. Students who take this approach are able to apply knowledge to new situations, understand text and produce written answers at a higher level than those who adopt a surface approach (Table 1.5) (Burge, 2003:243).

Table 1.5: Characteristics of students with different learning approaches (Burge, 2003:243)

Deep-learning approach • Intend to understand and actively seek meaning to satisfy curiosity • Understand the relationship between facts and concepts • Relate new ideas to their previous knowledge and personal experiences • Can analyse a professional situation and focus on the critical aspects • Question and are able to explain topics by reconstructing knowledge • Enjoy and are interested in their work • Are prepared to spend more time in independent study than those with a surface approach • Are motivated by an interest in the subject and / or recognition of relevance to vocation • Retain facts some weeks later

Table 1.5 - continue: Characteristics of students with different learning approaches (Burge, 2003:243)

Surface-learning approach
• Memorise facts for assessments without attempting to understand meaning • Accumulate unrelated facts and treat related parts separately • Reproduce essentials as accurately as possible • Show no evidence of refection on purpose or strategy • Find an answer to a problem without grasping the underlying issues or principles illustrated by the problem • Meet demands of task with minimum of effort • Are motivated by a desire to complete task or fear of failure

1.3 South African regulatory requirements

The Health Professionals Council of South Africa (HPCSA) oversees the quality control of undergraduate curricula. Guidelines laid down by the HPCSA are comprehensive, but not too restrictive, so that some freedom of choice is allowed for different schools. That is one of the main reasons why differences among medical schools' curricula exist (McKimm and Jollie, 2004:2).

The following summary from Part II (Framework of the undergraduate curriculum in medicine profile of the basic medical practitioner) of the Guidelines of the Health Professions Council of South Africa is important because it shows the need for a statistical and / or research methodology course to be part of the undergraduate training of medical students in South Africa (HPCSA, 1999:5, 9).

Students need to be able to develop their research abilities. The students must have the understanding of scientific principles but also be capable of making the right decisions (with the help of critical reading). One of the knowledge objectives is that the student must have an understanding of research methods (HPCSA, 1999:5, 9).

In South Africa, according to the South African Qualifications Authority (SAQA) regulations, on completion of an education and training programme, learners should be able to:

- Collect, analyse, organise and critically evaluate information.
- Communicate effectively, using visual, mathematical and/or language skills in the modes of oral and/or written presentations (Republic of South Africa, 1998: 3).

The Health Professions Council of South Africa implemented a compulsory programme of Continuing Professional Development (CPD) for doctors in 1999. The objective of the CPD programme is to improve patient care and simultaneously ensure that members of the medical profession maintain and improve their skills. CPD is an evolving programme, and various changes have occurred in the system since its introduction. The South African Medical Association (SAMA) fully supports this programme as a worthwhile development in the practice of medicine, and believes that all doctors should maintain professional ethics and the provision of quality health care throughout their careers. The Association is an accreditor for CPD activities and a member of the HPCSA Accreditors Forum that meets regularly to address issues pertaining to the programme. As an accreditor SAMA

- assesses CPD applications,
- responds to CPD queries,
- co-ordinates the CPD Committee,
- develops CPD policy and system (http://www.sama.co.za).

Continuing professional development is essential in a country characterised by an uneven geographic distribution of wealth and health resources, in which access to academic centres is much easier for those living in large cities (Lejarraga et al., 1998: 562).

1.4 Conclusion

Today it is easier to justify the place and importance of a medical statistics course than as recently as a decade ago and there exists widespread agreement on the necessity of medical students, to become at least consumers of research (Palmer, 2002:997).

We are training medical students to become mainly consumers, not producers, of research, and we must therefore not lose sight of the main aim of medical education: to produce better doctors, who deliver high quality health care (Campbell, 2002:4).

Medical doctors and health related professionals need to understand the process of statistical investigations and be able to plan statistical inquiry in medical and health related decisions (Bazargan, 2002:1).

If the aim of the medical statistics course is that students can understand and interpret statistical analyses reported in the literature, then it is important that this, is, in fact what is assessed in the course (Astin *et al*., 2002:1006).

The aims and objectives of the teaching of medical statistics have leaned towards the application of EBM. The content has altered as well, but the outcome of such a change remains unknown, but may include a threat to the professional identity of medical statisticians, and a decline in the understanding of traditional statistical concepts. If the objective is to promote EBM, we should take every opportunity to empower future doctors to critically appraise, and then apply, the results of well-conducted studies of important clinical questions (Morris, 2002:976).

In general, a medical student's focus is on the acquisition of skills needed to practice clinical medicine, and great care must be taken to explain why disciplines such as statistics are relevant to this. The use of real examples and an emphasis on the need for evidence has meant that medical students are increasingly aware of the pressure experienced by clinicians to justify their treatment decisions, and the associated need to be able to understand and critically appraise medical research (Astin *et al*., 2002:1003).

It is often assumed that training health professionals in evidence-based medicine reduces unacceptable variation in clinical practice, and leads to improved patient outcomes. This will only be true if the training improves knowledge and skills and these, in turn, are translated into improved clinical decision-making (Fritsche *et al*., 2002:1338).

1.5 Aims and objectives

Statistics is a difficult topic to teach and learn and there is ample evidence that its application is often faulty in medicine as well as in many other scientific disciplines. Errors include aspects of design, analysis and reporting and interpretation (Garcia-Bethou and Alcaraz, 2004: 1).

During the 1970s, some medical schools already realised how important it was to instil a critical attitude among students. By 1990, a consensus had emerged that concepts, rather than arithmetic skills, were required, and more schools embraced critical appraisal. By 2000, the critical appraisal of medical literature had become a central theme (Morris, 2002:972).

It is therefore important to do this study on the profile of research methodology and statistics training for undergraduate medical students at South African universities in terms of the following:

(1) What subjects (topics) are medical students taught?

(2) Who does the teaching?

(3) When is the learning programme / contact sessions taught during the medical students' curriculum?

(4) How is the learning programme / contact sessions taught to the students?

CHAPTER 2

RESEARCH DESIGN AND METHODOLOGY

2.1 Study Design

This study is an observational descriptive study that consists of a quantitative checklist and a questionnaire with quantitative and qualitative questions.

2.2 Sample

Only a small population of universities train medical students in South Africa. I therefore included the whole population of these eight universities that train medical students in this study. The eight universities are:

- University of Cape Town, Cape Town
- University of Stellenbosch, Tygerberg
- University of the Free State, Bloemfontein
- University of the Witwatersrand, Johannesburg
- Walter Sisulu University (Mthatha campus), Transkei
- University of Kwazulu Natal, Durban
- University of Pretoria, Pretoria
- University of Limpopo- Medunsa Campus, Garankuwa.

2.3 Questionnaire Design / Measurement

See Appendix A for the checklist regarding the topics taught. The following four sources were used to compile the checklist:

1. Cheatham, M.L. (2000). A structured curriculum for improved resident education in Statistics. The American Surgeon, 66, 585-588.
2. Dixon, R.A. (1994). Medical statistics: content and objectives of a core course for medical students. Medical Education, 28, 59 – 67.
3. Altman, D.G. (1991). Practical Statistics for Medical Research, London: Chapman and Hall.
4. Dawson-Saunders, B. and Trapp, R.G. (1994). Basic & Clinical Biostatistics, 2nd ed., United States of America: Prentice-Hall International.

For each topic, respondents were asked whether the topic was introduced to the students. If it was, the respondent was asked whether the students need to calculate (where possible) the concept and whether the students need to use this. Lastly the respondent was asked whether the students need to interpret the concept as well. (For example, are the students introduced to the t-test, if they are do they need to calculate it and use the concept in tests and do they need to interpret the t-test.)

The remaining research questions regarding when the teaching is done, who does the teaching and how the teaching is done and some additional probing questions were used to compile the questionnaire. See Appendix B for the questionnaire.

2.4 Pilot study

The reasons why a pilot study was undertaken were because:

- I wanted to determine the feasibility of the study,
- To detect flaws in the protocol and measurement instruments,
- And to get an indication of possible results (Joubert *et al.*, 1999:52).

Subjects in the pilot study should be as similar as possible to those in the study population, but must preferably not be part of the sample. Because the population is so small, it was decided to conduct the pilot study on the nursing learning programme at the University of North West (old University of Potchefstroom) and the dental learning programme at the University of Johannesburg (old Rand Afrikaans University). I received consent from the Head of the Department of Statistics at the University of Johannesburg to do my pilot study, but because of problems getting in touch with lecturers it was impossible to do the pilot study at the University of Johannesburg in the time frame available for me to complete my studies. I then contacted the Head of research at the School of Nursing Science at the University of the North West per email to find out whether they teach any statistics or research methodology courses, and if they would be willing to help me with my pilot study. The answer was positive and I contacted the relevant person. We discussed my study, and I forwarded the checklist and questionnaire to her. It was decided that we meet. I went to interview the contact person to complete the questionnaire and checklist and to find out if I needed to add anything else to the questionnaire and checklist. Out of this pilot study section 15c – d were included in the checklist and a choice between a few extra statistical software packages were included under question 9. Question 10 was also included in the questionnaire.

2.5 Data collection and fieldwork practice

I contacted the heads of the eight medical schools via email to ask them whether they would give consent for their university to participate in my study. If they were willing to include their university in my studies I asked them to complete the consent form (Appendix C - Afrikaans or Appendix D - English) and fax it back to me with the contact details of the person(s) that are responsible for the teaching of statistics and research methodology to the medical students. Two of the heads completed the consent form and returned it to me via fax, one sent me a formal letter confirming their participation. The rest of the heads of the medical schools gave their consent, not on their consent form, but indicating by email that they did not see any problems with my request, and that I could contact the relevant persons, as indicated in their emails. One university was not able to participate in the study, due to unforeseen circumstances. My contact person had fallen ill and was thus unable to assist me, and nobody else was available or had the necessary background of the subject matter to help me.

I contacted the relevant persons of all the universities via email and asked them if they were willing to participate in my study and whether they were willing to give me permission to interview them and complete a consent form (Appendix E). All the respondents gave me consent, but not on the consent form but via email. If they were willing to participate, I asked them if a personal interview would be possible. I visited two universities for the personal interview with the relevant contact persons, where we completed the questionnaire and checklist. Respondents for the other five universities did not have any problems with participating in the study, but were more eager to complete the questionnaire and checklist via email. Four of the respondents completed the questionnaire and checklist that I provided via email. For the other university the questionnaire and checklist could not be completed since no course is presented in statistics. They did give me some helpful feedback via email and telephone conversations, and this information I used in the completion of the questionnaire and checklist. One university did not participate in the study. I searched all the universities Internet websites, to see if there is anything that could add insight to my studies.

The least emails I sent out to get a response was three, and the largest number of emails necessary for me to send out until I received information was 35 emails, not including reminders. The average number of emails necessary to receive feedback was 14 emails. This did not include the reminders or any follow up emails, just emails trying to get in touch with the correct person to complete the questionnaire and checklist.

I also requested the learning programme material from the universities to see what the aims and objectives of their courses are. Five universities supplied me with their learning program materials.

2.6 Data Management

The interviewer recorded all the data directly onto standardized questionnaires and checklists for the personal interviews. The other respondents completed the checklist and questionnaire themselves and returned them by email. Then I created a sample checklist on a MS Excel spreadsheet. I captured all the information that I received from each of the completed checklist onto my sample checklist. Thus I entered seven completed checklists onto my sample checklist (each time increasing the incidence of each topic by one, when a respondent indicated that the topic was taught at their university). I also created a sample questionnaire on a MS Word document. I captured all the information that I received from each of the completed questionnaires onto my sample questionnaire. Thus I entered seven completed questionnaires onto my sample questionnaire (each time adding the new information received from the respondents on each question completed).

2.7 Analysis of the Data

Data checking was the first step in the data analysis process. The results are displayed in tables. The figures in the tables are the number of universities that teaches the specific concept. Thus for each section of the checklist a corresponding section will be in chapter 3 and chapter 4. The responses to the open questions were grouped into themes. These themes were tabled.

2.8 Response Rate

The response rate for the study is 87.5%, since information was obtained from 7 of the 8 universities. Seven out of the 8 universities supplied me information regarding my questionnaire and checklist. Five out of 8 universities supplied me information regarding the learning programme material.

2.9 Ethical Aspects

I obtained consent from the relevant Heads of the Schools of Medicine at the eight universities for the participation in the study and this was on a voluntary basis. The consent form was available in Afrikaans (Appendix C) and English (Appendix D). I received consent from all the respondents that is lecturers or course administrators (respondents' consent form in Appendix E). The subjects were informed that the study results would be given to them. Reporting about the study was honest, frank and sensitive. The sources of the information were kept confidential and results were reported anonymously. Feedback will be given to each respondent. The protocol was approved by the Ethics Committee of the Faculty of Health Sciences of the University of the Free State.

CHAPTER 3

RESULTS

3.1 Introduction

In this chapter I present the results of the checklist (Appendix A) and the questionnaire (Appendix B) that were completed for each participating university. I also include the summary of the aims and objectives of the learning programmes of the universities that submitted their learning programme material. The checklist is presented in section 3.2; the questionnaire is presented in section 3.3 and the learning programme material in section 3.4. The summary in section 3.5 ends off this chapter.

Five of the eight universities provide a formal Biostatistics / Research Methodology course. Of the remaining three universities one university did not participate in this study, one university teaches a few topics of Statistics / Research and at the third university, the research subject is not part of the curriculum, it is part of the electives programme. Therefore, those students who do not elect the research subject will not be exposed to statistics or research methodology. For the study results presented, the elective programme was handled as if all the students of that university were exposed to the concepts (because the topics are presented).

3.2 Checklist

The relevant person at each of the universities completed the checklist (Appendix A). The persons who completed the questionnaires are a doctor (3 universities), statistician (6 universities) and Applied Mathematics lecturer (1 university). At some universities more than one person completed the questionnaire and checklist and both will be indicated in the analysis if both teach the topic. That is why in some cases in sections 3.2.1 to 3.2.16 the number of different persons will not add up to seven, but sometimes more than that. The five universities that do a statistics / research methodology course, the university that does an elective program and the university that does some selective topics were included in the section. Thus a checklist was completed for all seven universities. One university does a repeat of the statistics and research methodology course in the 3rd and 5th year of the medical curriculum. This entails a brief summary of the topics covered. This is to refresh the medical students' statistical knowledge, to equip them to be able to critically appraise the medical literature that they are studying. For the university that does not do a statistics / research methodology course, they do a "Research" topic that is done in a time period of 4 weeks (in their 2nd year), in what is called a "block".

In Table 3.1 a summary is given of who teaches what topics at the different universities. It is clear that a statistician teaches most of the topics. The Public Health physician and Applied Mathematician are used at two universities only.

Table 3.1: Summary of who teaches what topics (n = 7)

Topic	Doctor	Statistician	Public Health Physician	Applied Mathematician
Introduction to medical research	3	3	1	1
Study designs in medical research	2	3	1	-
Exploring and presenting data	-	5	1	1
Summarising data	1	5	1	1
Probability	-	4	1	1
Sampling	-	5	1	1
Clinical measurement	1	2	1	-
From sample to population	-	4	-	1
Statistical inference	-	6	1	1
Analysis of the means of small samples	-	3	1	1
Analysis of cross tabulations	-	4	1	1
Methods based on rank order	-	1	1	-
Multiple comparisons	-	2	1	-
Correlation and regression	-	4	1	1
Other topic: Critical reading	1	4	1	1
Additional topics	2	3	1	-

In Table 3.2 a summary is given of when the topics are taught at the different universities. It is clear that most of the topics are taught during the 1st year of the medical curriculum.

Table 3.2: Summary of when the topics are taught (n = 7)

Topic	1st year	2nd year	3rd year	5th year	3rd / 4th year (during electives)
Introduction to medical research	5	1	1	1	1
Study designs in medical research	4	1	1	1	1
Exploring and presenting data	4	1	1	1	1
Summarising data	4	2	1	1	1
Probability	4	-	-	-	1
Sampling	4	1	-	-	1
Clinical measurement	3	-	-	-	-
From sample to population	3	-	-	-	1
Statistical inference	4	1	1	1	1
Analysis of the means of small samples	3	-	-	-	1
Analysis of cross tabulations	3	1	-	-	1
Methods based on rank order	2	-	-	-	-
Multiple comparisons	2	-	-	-	1
Correlation and regression	3	1	-	-	1
Other topic: Critical reading	5	1	1	1	1
Additional topics	5	1	-	-	-

3.2.1 Introduction to medical research

All the universities (7 universities) give a lecture on the introduction of medical research to the medical students.

3.2.2 Study Designs in Medical Research

Table 3.3: Presentation of topics regarding study designs in medical research (n = 7)

Topics	Introduce to concept	Need to calculate and use concept	Need to interpret the concept
	Yes	**Yes**	**Yes**
a. Observational studies			
1. Cross sectional surveys	5	2	4
2. Cohort studies	6	3	5
3. Case control studies	5	2	4
4. Case series	4	1	2
b. Experimental studies 1. Controlled trials			
1.1 Randomised controlled clinical trials	5	2	4
1.2 Not randomised	4	1	3
1.3 Sequential trials / or interim analysis	2	0	1
1.4 External controls	3	0	1
2. Uncontrolled trials	4	0	2
c. Association and causality	5	5	2
d. Retrospective vs. prospective	5	5	4
e. Longitudinal vs. prevalence studies	5	5	4

From Table 3.3 it can be seen that two universities teach all the controlled experimental studies and five universities introduce randomised controlled trials to the medical students. At five of the universities the concepts of association, causality, retrospective studies, prospective studies, longitudinal studies and prevalence studies are taught. Most of these concepts need to be interpreted but not used.

3.2.3 Exploring and presenting data

Table 3.4: Presentation of topics regarding exploring and presenting data (n = 7)

Topics	Introduce to concept	Need to calculate and use concept	Need to interpret the concept
	Yes	Yes	Yes
1. Nominal scales			
1.1 Dichotomous / binary	6	5	5
1.2 Qualitative / categorical observations	6	5	5
1.3 Percentages / proportions	6	5	5
2. Ordinal scales	6	5	5
3. Numerical scales			
3.1 Quantitative	6	5	5
3.2 Continuous	6	5	5
3.3 Discrete	6	5	5
4. Tables and graphs for nominal and ordinal data			
4.1 Frequency distribution	6	4	4
4.2 Bar charts	6	4	4
4.3 Contingency table	6	4	4

Table 3.4-continue: Presentation of topics regarding exploring and presenting data (n = 7)

Topics	Introduce to concept	Need to calculate and use concept	Need to interpret the concept
5. Tables and graphs for numerical data	**Yes**	**Yes**	**Yes**
5.1 Stem and leaf plots	4	4	3
5.2 Frequency tables	6	4	4
5.3 Histograms	6	4	4
5.4 Box and whisker plots	5	5	4
5.5 Pie charts	6	4	3

Table 3.4 indicates that six universities introduce nominal, ordinal and numerical scales. At 5 of the universities, students need to use and interpret the concepts (nominal, ordinal and numerical scales). Four of the respondents indicated that the universities introduce all the tables and graphs.

3.2.4 Summarising data

Table 3.5: Presentation of topics regarding summarising data (n = 7)

Topics	Introduce to concept	Need to calculate and use concept	Need to interpret the concept
	Yes	**Yes**	**Yes**
a. Measures of the middle (Central tendency)			
Mean	7	5	5
Weighted average	4	3	3
Median	7	5	5
Mode	7	4	4
Geometric mean	2	2	2

Table 3.5 - continue: Presentation of topics regarding summarising data (n = 7)

b. Measures of spread			
Range	7	4	5
Standard deviation	7	4	5
Variance	6	3	4
Coefficient of variation	6	3	4
Standard error	7	3	4
Percentiles	7	4	5
Interquartile range	6	4	5
c. Measures to use with nominal data			
Proportions and percentages	6	5	5
Ratios and rates	6	5	5
Vital statistics rates	4	4	3
1. Mortality rates	4	4	3
2. Morbidity rates	4	4	3
3. Prevalence	5	4	4
4. Incidence	5	4	4
5. Adjusted rates			
5.1 Direct method of adjusting rates	3	2	2
5.2 Indirect method of adjusting rates	2	1	1
d. Measures to describe relationships between two characteristics			
1. Describing the relationship between two characteristics			
1.1 Correlation coefficient	4	3	4
1.2 Coefficient of determination	3	2	3
2. Describing the relationship between two ordinal characteristics			
2.1 Spearman rank correlation	2	2	2
3. Describing the relationship between two nominal characteristics			
3.1 Relative risk	4	3	4
3.2 Odds ratio	4	3	4
e. Reliability	1	1	1

Table 3.5 shows that the basic central tendencies (mean, median and mode) are introduced at all participating universities. At 4 of the universities students need to interpret the concept, need to calculate and use these concepts. Six of the universities introduced the concept of "Measures of spread" to their medical students. At 3 of the universities the students need to calculate, use and interpret these concepts. At three of the universities students need to calculate, use and interpret the concepts of mortality and morbidity rates. At four of the universities students need to calculate, use and interpret the concepts of prevalence and incidence. At five of the universities students need to calculate, use and interpret the concepts of proportions, percentages, ratios and rates. Adjusted rates were a topic that was not covered by many universities. At three of the universities students need to calculate, use and interpret the concepts of correlation coefficient, relative risk and odds ratio. The rest of the topics under "Measures to describe relationships between two characteristics" and reliability were not covered by many universities.

3.2.5 Probability

Table 3.6: Presentation of topics regarding probability (n = 7)

Topics	Introduce to concept	Need to calculate and use concept	Need to interpret the concept
	Yes	Yes	Yes
Random variables & probability distributions			
1. Binomial distribution	4	2	4
2. Poisson distribution	2	2	2
3. Normal distribution	5	3	5

It is clear from Table 3.6 that 5 of the universities introduce the normal distribution to the medical students. At 5 universities the students need to interpret the concept. The other two distributions (Binominal and Poisson) are covered by fewer of the universities.

3.2.6 Sampling

Table 3.7: Presentation of topics regarding sampling (n = 7)

Topics	Introduce to concept	Need to calculate and use concept	Need to interpret the concept
	Yes	**Yes**	**Yes**
Random sampling	6	4	5
Systematic sampling	5	3	4
Stratified sampling	4	2	3
Cluster sampling	5	3	4
Non-probability sampling	4	2	3

From Table 3.7 it is clear that six universities teach random sampling (these concepts need to be interpreted and used by students at 4 of the universities), whereas fewer universities cover specific sampling methods.

3.2.7 Clinical Measurement

Table 3.8: Presentation of topics regarding Clinical Measurement (n = 7)

Topics	Introduce to concept	Need to calculate and use concept	Need to interpret the concept
	Yes	Yes	Yes
Repeatability and precision in measurements	3	1	2
Digit preference	0	0	0
Comparing two methods of measurement	3	1	2
Sensitivity and specificy	3	2	2
Positive predictive value	3	3	3
Negative predictive value	3	2	3
Accuracy	2	2	2
Receiver operating curves (ROC)	0	0	0
Survival data	1	1	1
Computer aided diagnosis	0	0	0
Kaplan-Meier curves	1	0	0
Life-table analysis	0	0	0
Actuarial analysis	0	0	0

Table 3.8 indicates that the topics under "Clinical measurement" are not introduced at a lot of the universities, even less of the universities' students need to use, calculate and interpret the concepts. Sensitivity, specificy, positive predictive value and negative predictive value are introduced at 3 universities to the medical students.

3.2.8 From Sample to Population

Table 3.9: Presentation of topics regarding from sample to population (n = 7)

Topics	Introduce to concept	Need to calculate and use concept	Need to interpret the concept
	Yes	Yes	Yes
Properties of the normal distribution	4	3	4
Reference ranges	3	3	3
Standard error of a sample mean	4	3	4
Standard error of a proportion	4	3	4
Standard error of the difference between two means	3	3	3
Standard error of the difference between two proportions	3	3	3
Sample size for an estimate	4	3	3

Table 3.9 shows that all the topics under "From sample to population" are introduced at three of the universities and the students need to use and interpret the concepts.

3.2.9 Statistical inference

Table 3.10: Presentation of topics regarding statistical inference (n = 7)

Topics	Introduce to concept	Need to calculate and use concept	Need to interpret the concept
	Yes	**Yes**	**Yes**
Testing a hypothesis	6	4	6
The null hypothesis	6	4	6
Type one/two errors	5	3	4
The p-value	6	4	6
Significance levels	6	4	6
One-and two-sided tests of significance	5	3	4
Significant, real and important	4	2	3
Statistical power	4	3	4
Degrees of freedom	5	4	5
Confidence intervals rather than p-values	5	3	4
Comparing the means of large samples using the normal distribution	3	2	2
Comparison of two proportions	2	1	1
Multiple significance tests	0	0	0
Exploratory data analysis	2	1	2

It is clear from Table 3.10 that the following topics under "Statistical inference"; "testing a hypothesis, the null hypothesis, the p-value and significance levels" are introduced and need to be interpreted by the medical students at 6 of the universities. The other concepts are not covered by many of the universities.

3.2.10 Analysis of the means of small samples

Table 3.11: Presentation of topics regarding analysis of the means of small samples (n = 7)

Topics	Introduce to concept	Need to calculate and use concept	Need to interpret the concept
	Yes	Yes	Yes
The t distribution	4	3	4
The one sample t method	4	3	4
The means of two independent samples	4	3	4

From Table 3.11 it follows that that 4 of the universities teach all the topics under "Analysis of the means of small samples" to the students. The students are introduced to the topics and they need to interpret the concept and need to calculate and use the concepts.

3.2.11 Analysis of cross tabulations

Table 3.12: Presentation of topics regarding the analysis of cross tabulations (n = 7)

Topics	Introduce to concept	Need to calculate and use concept	Need to interpret the concept
	Yes	Yes	Yes
The chi-squared test for association	5	4	5
Validity of the chi-squared test for small samples	2	2	2
Fisher's exact test	4	3	4
Yates` continuity correction for a 2 by 2 table	2	2	2
McNemar`s test for matched samples	2	0	1

Table 3.12 reveals that 5 universities teach the topic of “chi-squared test” (the students are introduced to the topics and they need to interpret the concept and need to calculate and use the concepts) and the rest of the topics are taught at fewer universities.

3.2.12 Methods based on rank order

Table 3.13: Presentation of topics regarding the methods based on rank order (n = 7)

Topics	Introduce to concept	Need to calculate and use concept	Need to interpret the concept
	Yes	Yes	Yes
The Mann-Whitney U test	2	1	2
Wilcoxon matched pairs test	2	1	2
Kendall’s rank correlation coefficient	0	0	0

Table 3.13 shows that only two universities introduce the topics of “the Mann-Whitney U test” and the “Wilcoxon matched pairs test” (two universities’ students need to interpret these concepts, but only one of the universities’ students need to calculate and use the concepts).

3.2.13 Multiple comparisons

Table 3.14: Presentation of topics regarding multiple comparisons (n = 7)

Topics	Introduce to concept	Need to calculate and use concept	Need to interpret the concept
	Yes	Yes	Yes
ANOVA	3	2	3
Bonferroni adjustments	1	1	1
Meta-analysis	1	0	0
Post Hoc comparisons	0	0	0

It is clear from Table 3.14 that three universities teach ANOVA (this concept is introduced to the students and need to be interpreted by the students) whereas fewer or no universities cover the other topics.

3.2.14 Correlation and regression

Table 3.15: Presentation of topics regarding correlation and regression (n = 7)

Topics	Introduce to concept Yes	Need to calculate and use concept Yes	Need to interpret the concept Yes
Scatter diagrams	4	2	4
Confidence interval and / or significance test for the correlation coefficient	4	1	2
The method of least squares	2	2	2
Regression	3	2	2
Comparison of assumptions between correlation and regression	1	1	1
Multiple regression	2	1	1
Logistic regression	1	1	1

As can be seen from Table 3.15, four of the universities introduce scatter diagrams and confidence intervals. Only 3 universities introduce the topic of regression and at only 2 universities the students need to use, calculate and interpret these concepts. Some topics are introduced at even less universities.

3.2.15 Other topics

Table 3.16: Presentation of topics regarding other topics (n = 7)

Topics	Introduce to concept	Need to calculate and use concept	Need to interpret the concept
	Yes	Yes	Yes
Choosing the statistical method	4	2	3
Critical reading	7	5	6
Quasi - experimental designs	1	1	1
Quantitative designs	1	1	1

From Table 3.16 it is clear that seven of the universities introduce "critical reading". For five of the universities this topic needs to be interpreted and used, whereas fewer universities cover the rest of the topics.

3.2.16 Additional topics

The following are additional topics that the universities indicated they provide: The topics are computer literacy (1 university), qualitative studies (1 university), ethics (2 universities), protocol development (3 universities), questionnaire design (3 universities), trigraphs (1 university), truncated binomial distributions (1 university), average shifted histograms (1 university), normal probability plots (1 university), likelihood ratio (1 university), ecological study design (1 university), field trials (2 universities), field work (1 university), piloting (1 university), report writing (2 universities), Web CT (1 university), F-distribution (1 university), bias (1 university), evidence based medicine (1 university), data searches on the computer (1 university), univariate vs. multivariate analysis (1 university), multiple discriminant analysis (1 university), factor analysis (1 university) and cluster analysis (1 university).

3.3 Questionnaire

The relevant person at each of the universities completed the questionnaire (Appendix B). The persons who completed the questionnaires are a doctor (2 universities), statistician (6 universities) and Applied Mathematics lecturer (1 university). Only the five universities that do a statistics / research methodology course and the university that does an elective program was included in the section. Thus we only look at 6 universities when compiling the questionnaire.

The specific statistics / research methodology courses show a vast variety of implementation at the different universities. Two universities changed their course as little as 2 years ago (2003), two have it since the 1970/80's and two universities since 2000.

The contact sessions of the universities also vary from the shortest of 40 minutes to the longest of 180 minutes, with an average of 83 minutes. The number of contact sessions also varies. The two universities that have 40 and 45 minute long contact sessions have these contact sessions 6 times and 3 times per week for a whole semester. The other four universities that have 60 (2 universities), 90 and 180 minute contact sessions have 20, 5, 8 and 12 contact sessions.

The class sizes are very similar for 3 universities where the average size is around 200. One university has around 140 students, one around 320 students, and one 150 students (but the attendance will be around 40 students because this is part of the elective programme).

Table 3.17: Number of universities that does groupwork, practical classes, uses tutors and does a research project (n = 6)

	Yes	No
Do students do groupwork?	3	3
Do you use practical classes?	4	2
Do you use tutors?	3	3
Do students do a research project?	3	3

As indicated in Table 3.17 group work is done at three universities. At one university it forms part of the formal contact sessions but after the formal class (lecture), there is a group work session. At the second university the groupwork is done informally during tutorials. Students are encouraged to work in groups. They are supposed to discuss their answers of their tutorial work and ask any related questions about class work and tutorial work (what they do and do not understand). At the third university the groupwork is where the students must work on a project together and between 4 and 6 students work on such a project.

At four universities practical classes are used: three use the practical classes mainly for computer based problem solving and one as a platform for group work.

Three universities use tutors: two of the three universities use Honours in Statistics and Masters in Statistics students as their tutors, while the other university uses public health physicians, registrars and researchers as their tutors.

Table 3.17 shows that at three universities students have to do research projects. In the next paragraphs how this is done at each university is described.

University A

A topic is negotiated with community stakeholder. The students are randomised into groups and community based research is done. Protocol development, literature review, methodology take place. The next steps are fieldwork, analysis and report writing. Presentation takes place. Intervention such as some health promotion will take place.

University B

In the 2nd year the students are divided in groups of 4. The students are placed within a department with a study leader. In the 2nd semester of the 2nd year they must write a protocol. This is followed by a literature review, methodology, fieldwork, analysis and finishing of the report. In their 1st semester of their 3rd year they present their project.

University C

The students identify a topic they are interested in and a supervisor. They work in groups of about 4 students. They often join up with an existing study, as there is not enough time to develop a protocol, go through ethics and collect data themselves. The students have to analyse, write up a report and present the results.

In Table 3.18 is a list of problems that the universities encountered in the research methodology learning programmes or statistics learning programmes.

Table 3.18: Problems encountered in the statistical / research course (n = 6)

Repeating concepts (1 university).
Sometimes difficult to obtain appropriate data (mainly epidemiological data) (1 university).
Students are not much interested in statistics (2 universities). Problem is that students do not know why this course is necessary (1 university).
Lack of time to present complex subjects necessitates very brief and inadequate course (1 university).
Race quota system with selection makes it more difficult (1 university).
Didactic PowerPoint presentations (1 university).
Students do not come prepared to the tutorial sessions. They expect the tutors to do the work for them, or they do the tutorial work during the session instead of knowing what their problem areas are (1 university).
There is a wide range of mathematical ability and also some students have done Additional Maths and have some statistics while others have never done any stats – it takes the non-statistics people a lot of hard work to grasp the concepts and keep up (1 university).
Research projects are very time consuming, because sometimes there are nearly 35 projects a year (1 university).
Mathematical background is sometimes not up to par (1 university).
Lack of basic numeracy in students (1 university).
Some students have a problem with the control of the English and Afrikaans language (2 universities).
Diversity of students comes from different backgrounds and has different academic skills (1 university).

Out of Table 3.18, four basic areas of concern can be identified. The problem areas are (1) students, (2) data, (3) presentation and (4) time. The problems surrounding students are that they are not really interested in the subject.

Five of the universities expose the students to computer programs, directly in practical classes and indirectly through the research projects that the students must do. In Table 3.19 is a list of what computer programmes the students are exposed to and at how many universities. Three of the universities do not give any exposure to any statistical software package and one university does not give any exposure to any computer program (not Excel or any statistical software package).

Table 3.19: Computer programs students are exposed to (n = 6)

Program	Number of Universities
Excel	5
Statistica	1
Stata	1
EpiInfo	1
SPSS	1
Nvino	1

The views of what the respondents think the students on the research methodology / statistics learning programme according to the respondents are listed in Table 3.20 below.

Table 3.20: Respondents' opinions regarding student views on the research methodology / statistics learning programme (n = 6)

Quite positive (3 universities).
They find it challenging (1 university).
Some see the relevancy of the course and some think it is a waste of time but I do believe that the majority see the relevancy (1 university). Students see it as a compulsory module (1 university).
They view statistics as integrated part in Public Health teaching (1 university).
In their first year they are positive on the presentation but not so sure why they must do it. In their 2nd / 3rd year when they do their research project they are positive on the organisation of the course, and a better understanding on how it fits into the whole picture, but still some doubt (1 university).
English classes more positive about the course than the Afrikaans classes (1 university).
There are some who enjoy it and do very well if they identify the right project and supervisor (1 university). Some students leave the project until the last moment and then they battle (1 university).

Interesting to note from Table 3.20 there are the two opposites: one group of students that find the course quite positive, and another group of students that does not see the necessity of the course.

As one can see from Table 3.21, four universities teach the statistics and research methodology modules as an integrated module, one university as a separate module and one university does not teach a research methodology course. This means that four universities teach the research methodology and statistics as one course, with no distinction between the subjects. One university teaches the research methodology and statistics course as two separate courses and one university does not teach any research methodology.

Table 3.21: Are the statistics and research methodology taught as separate modules or an integrated module (n = 6)?

Separate Module	**Integrated Module**	**No Research Methodology**
1	4	1

The stages when the theory of the statistics is used in other modules are listed in Table 3.22.

Table 3.22: Stages of implementation of statistics in other modules (n = 6)

In the 4th year => There are academic afternoons when they look at EBM (1 university).
During 4th / 5th year => when they have classes with some doctors (1 university).
In the special study modules (SSM) (1 university).
In research methodology module (1 university).
No stages (2 universities).
It is a golden thread and every clinical module must revisit the biostatistics concepts – it is a policy of the medical school (1 university).
Do not know (1 university).

At three universities the respondents did not know whether the statistics content is implemented again or knew that it is not implemented at all. At the other three universities the statistics content is implemented across the medical curricula, at some universities more that once.

Some suggestions by the respondents on improvements of the course are listed in Table 3.23. Three of the suggestions surround timing, the other suggestions are, more practical examples and more critical reading.

Table 3.23: Suggestions of improvements (n = 6)

Research projects that run during 2nd / 3rd year, to run simultaneously with 1st year. (1 university)
Applications made more applicable, with more critical reading. (1 university)
If it could be later during the medical studies (instead of 1st year) they can have the clinical knowledge. (1 university)
More time allocated to module. (2 universities)
More practical examples (and more new up to date examples) for the tutorials. The students should actually see when and where to use the application, which were taught to them. (2 universities)
Tutorials and tests must move to Web CT, which is an internet driven software to provide computer aided instruction to students to help them to work through examples on different topics, which provide them with answers to every multiples choice question even when they clicked on the wrong answer it helps them to understand why they made the wrong decision. (1 university)

3.4 Learning Programme Material

I received the learning programme material of 5 of the medical schools. I will briefly give a summary of the main aims and objectives of the learning programme as stated in the learning programme material.

The main aims and objectives of the learning programmes of the five universities are:

University A: To develop intellectually, into clear and independent thinkers who can make informed decisions and provide leadership. These would encompass the following: (1) analytical and critical thinking skills, (2) problem-solving, (3) numeracy, (4) computer literacy, (5) basic understanding of research methods.

University B: The module aims to provide an introductory course of the basics of Epidemiology and Biostatistics in healthcare. The module aims to expose students to all the steps involved in doing medical research. The student should gain the research skills, which are essential for their future career.

University C – elective course: The Biostatistics course is an introduction to research methodology and statistical analysis, because not necessary to do any statistical tests, but to understand the logic of statistical thinking, and which tests to use when.

University D: The course serves as an introduction to some areas of statistics and computing that have relevance in the biomedical field. The objectives of the course are: (1) to give an overview of techniques and processes involved in statistical experiments, (2) to enable students to perform the basic quantitative methods of biomedical research to analyse data, (3) to introduce the use of statistical software packages in analysing data, (4) to enable students to read and understand biomedical literature in which statistical concepts are used.

University E: This module will help the student to become an informed user of medical data. The following need to be achieved by the students: (1) Develop good judgement when it comes to clinical data, (2) critical reading, and understanding of mistakes in literature, (3) interpret the important survival statistics to use in epidemiology, (4) make the correct deductions from diagnostic procedures and laboratory test results, (5) interpret the results of the manufacturing of medicine and instruments, (6) evaluate study protocols and articles that are presented for publication in medical journals.

3.5 Summary

The following is clear from section 3.2 in this chapter. All the universities' students are introduced to medical research. The basic summarising data tools (mean, median, mode, range, standard deviation and standard error) are introduced at seven of the universities. The other summarising data tools are taught at fewer of the universities. Testing hypothesis, the null hypothesis, the p-value and significance levels are taught at 6 universities. The other topics under statistical inference are taught at fewer universities. At seven universities the medical students are introduced to the topic of critical reading.

Most of the topics under exploring and presenting data are introduced to six of the universities' students. Random sampling is introduced at 6 of the universities to the medical students, and the other sampling topics at fewer universities. The more common study designs (cross sectional, cohort, case control and case series) are introduced to 5 of the universities' students. The other study designs are taught at fewer universities. Five universities introduce the topic of the normal distribution to the medical students, and fewer universities cover the rest of the probability topics. At five universities the topic of chi-squared are introduced to the medical students. The other topics under "Analysis of cross tabulations" are covered by fewer of the universities. The t-distribution is taught at 4 universities to the medical students.

Clinical measurements are not a very popular topic for teaching and very few of the universities teach this to the students. Only 3 universities teach the topics of sensitivity, specificy, positive predictive value and negative predictive value to medical students. Only a few universities cover the topics under “methods based on rank order”. Multiple comparisons, correlation and regression are topics covered by only a few universities.

The following is clear from section 3.3 in this chapter. The statistics / research methodology course has been changed in the last years for a few universities, while for other the course was not changed for the last 20 to 30 years. The contact sessions vary from 40 minutes to 180 minutes, and from as little as five sessions to 6 times a week for a whole semester. The class sizes vary from 40 to 320, with an average of 200.

At three universities the students do groupwork, and a research project. At 3 universities the help of tutors are used, and at 4 universities students need to attend practical classes. The problems encountered with the research methodology / statistics course are mainly lack of skills and knowledge of the students to master this course and the lack of knowledge of the necessity of the course. Time constraints are also a major problem. Five universities expose the students to Excel, three of the universities give exposure to statistical software packages and one university does not give any exposure to any computer program.

Four universities teach the statistics and research methodology modules as an integrated module, one university as a separate module and one university does not teach a research methodology course. Three universities' respondents did not know whether the statistics content is implemented again or knew that it is not implemented at all. At the other universities it is implemented, at some universities more that once.

In the next chapter the results from the questionnaire and checklist will be discussed. I will analyse the results in conjunction with what is proposed in the literature.

CHAPTER 4

DISCUSSION

4.1 Introduction

In this chapter I discuss the results of the checklist (Appendix A) and the questionnaire (Appendix B) that were completed for each university in context with the literature. The checklist is presented in section 4.2, the questionnaire in section 4.3, the learning programme material in section 4.4, the conclusion in section 4.5 and the shortcomings in section 4.6 ends off this chapter.

Of the seven participating universities five provide a formal Biostatistics/ Research Methodology course, one university teaches a few topics of Statistics / Research and one presents the research subject as part of the electives programme.

4.2 Checklist

4.2.1 Study Designs in Medical Research

The extent to which the study design concepts are introduced at South African universities differ. Five universities' students are introduced to the topics of case control and cross sectional studies, and at 6 universities' students are introduced to cohort studies. At five universities randomised controlled clinical trials and at 4 universities not randomised clinical trials are introduced to the medical students. The rest of the experimental study topics are introduced at fewer universities to the medical students. Uncontrolled trials are introduced to 4 of the universities' students. The other concepts under experimental studies are taught at fewer universities. Association, causality, retrospective studies, prospective studies, longitudinal studies and prevalence studies are introduced at five universities to the medical students.

An important point which necessitates medical students to learn study designs was pointed out in 1990 by Appleton, namely that "It is unethical to carry out poorly designed studies on human subjects" (Appleton, 1990:1013). According to Sprent it is important that sound experimental design is used during your research in the laboratory (Sprent, 2003: 522).

According to Dixon the following are critical under design of studies:

- explain and apply the principles of randomised controlled clinical trials
- explain the difference between association and causality,

and these form part of the core curriculum used at the University of Sheffield, United Kingdom (Dixon, 1994:64).

Cheatham also suggests that the topics of study design need to be part of a medical statistics course because one needs to equip medical students with the statistical knowledge to perform the correct and necessary research (Cheatham, 2000:586).

According to a study done by Looney *et al.*, the medical schools in the United States that have a Biostatistics course cover the following topics: 91% cover case-control studies, 91% cover cohort studies, 88% cover randomised clinical trials, 88% cover study design characteristics, 81% cover descriptive studies and 77% cover cross sectional studies (Looney *et al.*, 1998:93).

One of medical research's biggest problems of error includes the aspect of design and it is critical to teach the medical students to know which study design they need to use (Garcia-Bethou and Alcaraz, 2004: 1).

Different types of studies are used in medical research. Knowing how a study is designed is important for an understanding of the conclusions that can be drawn from that (Dawson-Saunders and Trapp, 1994:6).

It seems that what the South African universities introduce to the medical students regarding study design is mainly sufficient and in line with what is expected of medical students in other medical curricula in other parts of the world. I will recommend that all the universities need to introduce all the types of study designs because the subject of study design is such an important one.

4.2.2 Exploring and presenting data

Six universities in South Africa indicated that they do nominal scales of measurement (binary, categorical, proportions), ordinal scales of measurement and numerical scales of measurement (quantitative, continuous, discrete). All the above mentioned topics need to be calculated, used and interpreted by the students at five universities. Six universities introduce the concepts of frequency distribution, bar charts and contingency tables under the topic of tables and graphs for nominal and ordinal data. Under the topic of tables and graphs for numerical data, six universities introduce the frequency tables, histograms and pie charts, whereas 5 universities introduce box and whisker-plots and 4 universities introduce stem-and leaf-plots to the medical students.

Cheatham suggests that the following topics on the types of data are the most important topics: binary data variables and continuous data variables (Cheatham, 2000: 587).

According to Dixon the following is critical under presenting data:

- identify appropriate methods for presenting and summarizing clinical measurements (types of data, frequency distributions and histograms),

and this forms part of the core curriculum used at the University of Sheffield, United Kingdom (Dixon, 1994:64).

According to a study done by Looney *et al.*, the medical schools in the United States that have a Biostatistics course cover the following topics: 90% of the schools covers frequency distributions, 75% covers the interpretation of tables and graphs, 72% covers scales of measurement and 69% covers measurement issues (Looney *et al.*, 1998:93).

According to all the above examples the data presentation concept is very important for medical students. It teaches them to understand and interpret graph and data presentations that they will see in medical literature. One of the aims of teaching statistics to medical students according to the General Medical Council (GMC) is to educate doctors to be competent to interpret data presented in the press, pharmaceutical literature and professional journals (General Medical Council, 2003).

It is clear that the South African medical students are equipped with the knowledge that is necessary to be competent to interpret data.

4.2.3 Summarising data

Seven universities teach averages (means), median, mode, range, standard deviations, standard errors, percentiles, proportions and percentages and ratios and rates. At six universities the students are taught the concepts of variance, coefficient of variation, Interquartile range, prevalence and incidence. Five universities teach the concept of weighted average, mortality rates, morbidity rates, correlation coefficients, relative risk and odds ratio. All the other topics under “summarising data” are covered by fewer of the universities.

A popular topic in the Australian medical education is the average (mean) (Watson, 2002: 2). According to Simpson at the University of Sydney, descriptive statistics is a must for a medical statistics course (Simpson, 1995: 202). According to Sprent at the University of Dundee the students (or at a later stage the doctors) need to compute and quote standard deviations, standard errors and correlation coefficients (Sprent, 2003:522). One of the video topics at the University of Sheffield in the United Kingdom is on describing data (including all the summary measurements). This is given to the first year students (Campbell, 2002:4). One of the categories taught at the University of Tehran, Iran, is the term "survey method". In the survey research method, distribution of variable(s) under study is considered, and exploratory approach to understanding a problem is discussed (Bazargan, 2002:3).

Cheatham suggests that the following topics on descriptive statistics are the most important topics: the topics of the mean, median and mode, standard deviation and prevalence (Cheatham, 2000: 587).

According to Dixon the following is critical under summarising data:

- identify appropriate methods for presenting and summarizing clinical measurements (median, quantiles and / or percentiles, mean, variance and standard deviation),

and this forms part of the core curriculum used at the University of Sheffield, United Kingdom (Dixon, 1994:64).

According to a study done by Looney *et al.*, the medical schools in the United States that have a Biostatistics course cover the following topics: 92% of the schools covers descriptive statistics, 91% covers incidence and prevalence, 90% covers rates, 88% covers central tendency, 85% covers correlation and 64% covers adjusted rates (Looney *et al.*, 1998:93).

The reason why the summarising of data is important in a medical statistics course is because the students need to know what it means when they are confronted with the terms in medical literature. It is not necessary for them to know exactly how each of the summary tools needs to be calculated because that will be done by a statistical computer software package. It is also important to teach the topics of incidence and prevalence because these two concepts provide the basic description of disease occurrence. As long as they can distinguish between the different descriptive statistics and to know when to use the different statistical tools, they will master the medical literature (Astin *et al*, 2002: 1003).

It seems that what the South African universities introduce in terms of summarising data to the medical students is sufficient and is in line with what is expected of medical students in other medical curricula in other parts of the world. Only 3 of the universities introduce incidence and prevalence to the medical students in South Africa, and these topics need to be introduced at all the universities in South Africa.

4.2.4 Probability

The spectrum of the probability concepts that are introduced at South African universities differs. Five universities introduce normal probability distribution, whereas the rest of the probability concepts of binomial and Poisson distribution are introduced at fewer of the universities.

The terminology of probability (chance) is one of the popular topics in the Australian medical education (Watson, 2002: 2). According to Cheatham the topic of the normal distribution is important (Cheatham, 2000: 587).

According to Dixon the following are critical under probability:

- state the main properties of the normal distribution
- use tables of the standardized normal distribution,

and these form part of the core curriculum used at the University of Sheffield, United Kingdom (Dixon, 1994:64).

According to a study done by Looney *et al.*, the medical schools in the United States that have a Biostatistics course cover the following topics: 88% of the schools covers normal distribution, 78% covers probability and 49% covers binomial distributions (Looney *et al.*, 1998:93).

Probability concepts are helpful for understanding and interpreting data presented in tables and graphs in published articles. In addition, the concept of probability lets us make statements about how much confidence we have in estimates, such as means, proportions, or relative risks. Understanding probability is essential for the understanding the meaning of p-values given in journal articles (Dawson-Saunders and Trapp, 1994:66).

It seems that what the South African universities introduce in terms of probability to the medical students is sufficient and in line with what is expected of medical students in other medical curricula in other parts of the world. A worrying aspect is that not all the universities indicated that they introduce the students to the concept of normal distributions, and this concept needs to be introduced at all the universities in South Africa.

A shortcoming in the checklist is that the question was never ask if the students were introduced to the term probability and not just the probability distributions.

4.2.5 Sampling

The continuum of the sampling concepts that are introduced at South African universities differ. Six universities taught concepts of random sampling, whereas the rest of the sampling concepts are taught at fewer of the universities, systematic sampling (5 universities), cluster sampling (5 universities), stratified sampling (4 universities).

One of the most important topics in the Australian medical education are the understanding of sampling as indicated in the following extract from Watson "the dual notions of sampling and of making inferences about populations, based on samples, are fundamental to prediction and decision making in many aspects of life. Students will need a great many experiences to enable them to understand principles underlying sampling and statistical inference "

(Watson, 2002: 2). One of the video topics at the University of Sheffield in the United Kingdom covers sampling and confidence intervals and the idea of standard error. This is given to the first year students. It seems that students need to understand the concept of sampling so that they can have a better understanding and knowledge of statistical inference (Campbell, 2002:4).

Cheatham, Dixon and Looney *et al.* did not indicate the importance of the sampling topics. There are many reasons why medical students need to study sampling: samples can be studied more quickly than populations; a study of a sample is less expensive than a study of an entire population; a study of an entire population is impossible in most situations; sample results are often more accurate than results based on a population; if samples are properly selected, probability methods can be used to estimate the error in the resulting statistics; samples can be selected to reduce heterogeneity (Dawson-Saunders and Trapp, 1994:69-70).

It seems that what the South African universities introduce in terms of sampling to the medical students is sufficient and more than what is expected of medical students in other medical curricula in other parts of the world. More focus can, however, be placed on the different types of sampling techniques.

4.2.6 Clinical Measurement

Only a few of the clinical measurements topics are taught at South African medical schools. The topics of sensitivity, specificity, positive and negative predictive values, repeatability and precision in measurements and comparing two methods of measurement are taught at 3 universities.

The topics of sensitivity and specificity are taught in the medical statistics course at the University of Sydney, Australia (Simpson, 1995:202). At the School of Medicine, University of California, San Diego the following topics feature in the lecture content: sensitivity, specificity, positive predictive value and negative predictive value (Deutsch, 2002:3). Students at the Royal Free & University College Medical School are introduced to the topics of likelihood, prevalence and predictive values (Morris, 2002: 974).

Cheatham suggests that the following topics on epidemiologic tests are the most important namely sensitivity and accuracy (Cheatham, 2000: 587).

According to Dixon the following is part of a 12-topic course, namely; repeatability and precision in measurements, comparing two methods of measurement, sensitivity and specificity, survival data, but not critical enough to form part of the core curriculum used at the University of Sheffield, United Kingdom (Dixon, 1994:64).

According to a study done by Looney *et al.*, the medical schools in the United States that have a Biostatistics course cover the following topics: 77% of the schools covers characteristics of diagnostic tests and 77% covers interpretation of diagnostic tests (Looney *et al.*, 1998:93).

Although not seen as one of the core topics in the medical curricula, topics such as survival analysis will be encountered very often in the literature. As the topics of sensitivity and specificity are only introduced at three universities in South Africa I will recommend that these two topics need to be introduced at the other universities as well.

4.2.7 From Sample to Population

The topic of sample size (sample size for an estimate) is introduced at four universities in South Africa, as well as the concepts of properties of the normal distribution, standard error of a sample mean, standard error of a proportion. The rest of the concepts (reference ranges, standard error of the difference between two means and standard error of the difference between two proportions) are taught at fewer universities.

According to Dixon the following are critical under sampling and population:

- contrast the general meaning of the standard deviation and its specific use in a normal distribution (reference ranges)
- explain sampling variation (standard error of a sample mean, confidence intervals)

and these form part of the core curriculum used at the University of Sheffield, United Kingdom (Dixon, 1994:64).

Four of the universities in South Africa teach these topics, so it can be introduced to the other universities. It will be useful for the students to be taught about standard error in general, and not only in relation to sample means and proportions.

4.2.8 Statistical inference

Six universities teach the basic principles of statistical inference. These basic principles are, testing a hypothesis, null hypothesis, p-values and significance levels. Five of the universities teach the topics of Type I and Type II errors, one- and two-sided tests of significance, degrees of freedom and confidence intervals rather than p-values to the students. Four universities teach the topics of "significant, real and important" and "statistical power". The rest of the concepts of "comparing the means of large samples using the normal distribution", comparison of two proportions, multiple significance tests and exploratory data analysis are taught at fewer universities.

According to Sprent (at the University of Dundee, Scotland), p-values are important for the medical students to understand and are all too often misunderstood (Sprent, 2003:522). One of the supplement video topics at the University of Sheffield in the United Kingdom covers p-values, and accordingly there students should be able to understand the reasoning that leads to p-values and their limitations, but not be able to carry out chi-squared tests. This is given to the first year students (Campbell, 2002:4). The topic of the p-value is also taught in the medical statistics course at the University of Sydney, Australia (Simpson, 1995:202). Sterne suggests that one must consider the confidence interval along with the p-value to protect us from making mistakes especially when we are working with small samples. Along with this Sterne also suggests the teaching of the null hypothesis, the alternative hypothesis and the statistical power (Sterne, 2002: 993). At the University of Cambridge, United Kingdom, the teaching of the p-value is very important, but it is only taught in conjunction with many warnings concerning the methodology's potential and actual misuse, abuse and over-use. Terminologies like "null hypotheses, alternative hypotheses, Type I and Type II errors, power and statistically significant" need to be covered too, to equip students for reading the medical literature knowledgeably (Palmer, 2002:996). The concept of testing of the hypothesis for two sample means are introduced and discussed to the students at the University of Tehran, Iran (Bazargan, 2002:3). At the University of California, the confidence interval is introduced and the application explained the medical students (Deutsch, 2002:3).

In a 12-month course recommended by Cheatham the following basic statistical theory are part of that course; hypothesis testing, type I / II errors, significance levels, power, one-tailed vs. two-tailed testing, and all of the topics are important (Cheatham, 2000: 587).

According to Dixon the following are critical under statistical inference:

- explain statistical significance testing through an example based on the normal distribution (testing a hypothesis, the null hypothesis, the p-value)
- contrast the uses of the statistical deviation and the standard error of the mean (significance levels)

and these form part of the core curriculum used at the University of Sheffield, United Kingdom (Dixon, 1994:64).

According to a study done by Looney *et al.*, the medical schools in the United States that have a Biostatistics course cover the following topics: 95% of the schools covers p-values, 93% covers hypothesis testing, 93% covers interpretation of confidence intervals, 82% covers construction of confidence intervals and 59% covers power analysis (Looney *et al.*, 1998:93).

It seems that what the South African universities introduce in terms of statistical inference to the medical students is sufficient and in line with what is expected of medical students in other medical curricula in other parts of the world.

4.2.9 Analysis of the means of small samples

Four universities in South Africa teach the concepts of the t-distribution, the one sample t method and the means of two independent samples.

According to Dixon the following is critical under one-sample t-test:

- appropriately apply the Student's paired t-test (degrees of freedom, the t distribution, the one-sample t-method)

and this forms part of the core curriculum used at the University of Sheffield, United Kingdom (Dixon, 1994:64).

It is clear from the literature that the topics under "Analysis of the means of small samples" are important. Four of the universities in South Africa teach these topics, so for the others that do not teach it to the students it would be recommended that they introduce these topics to the medical students.

4.2.10 Analysis of cross tabulations

Five universities in South Africa teach the topic of the chi-squared test for association. Four universities teach the topic of Fisher's exact test to the students. Fewer universities covered the rest of the topics (validity of the chi-squared test for small samples, Yates' continuity correction for a 2x2 table, McNemar`s test for matched samples).

According to Dixon the following are critical under chi square test:

- construct a two-way contingency table and formulate the appropriate null hypothesis (comparison of two proportions)
- apply the chi square formula to test for statistical significance (the chi squared test for association, validity of the chi squared test for small samples, standard error of the difference between two proportions)
- appropriately apply a formula for calculating the confidence interval for a proportion (standard error of a proportion)

and these form part of the core curriculum used at the University of Sheffield, United Kingdom (Dixon, 1994:64).

According to a study done by Looney *et al.*, 88% of the medical schools in the United States, which have a Biostatistics course cover, the chi-squared test (Looney *et al.*, 1998:93).

The chi-square test is the most commonly used method for comparing frequency or proportions, because it can be used with two or more groups. Chi-square also has a variety of other applications and is thus a very versatile test (Dawson-Saunders and Trapp, 1994:149).

It seems that what the South African universities introduce in terms of "Analysis of cross tabulations" to the medical students is sufficient and in line with what is expected of medical students in other medical curricula in other parts of the world. Two universities do not introduce this topic to the students and I would recommend that they do.

4.2.11 Methods based on rank order

The topics under methods based on rank order are done only at a few universities in South Africa. The topic of the Mann-Whitney U test and Wilcoxon matched pairs test are taught at two universities. No university in South Africa indicated that they introduced the topic of Kendall's rank correlation coefficient to the medical students.

According to a study done by Looney *et al.* only 29% of the medical schools in the United States which have a Biostatistics course cover the Wilcoxon Mann-Whitney test (Looney *et al.*, 1998:93).

Not Dixon or Cheatham says something specific about this topic. This is an important topic as one comes across a lot of skew data in medical research, and it is thus necessary to be introduced to these concepts.

4.2.12 Multiple comparisons

Three universities in South Africa introduce the concepts of ANOVA and one university the concept of Bonferroni adjustments. Only one university indicated that they teach the concept of meta-analysis.

According to the Nuffield Department of Clinical Medicine, Oxford, clinicians may make greater use of meta-analyses in the future; the ability to appraise critically publications of all types will remain an invaluable skill (Rosenberg and Donald, 1995: 1123). Sterne also stresses the importance of meta-analyses to make medical students better doctors one day (Sterne, 2002: 993).

Cheatham suggests that the following topics are important, the meta-analysis and Bonferroni adjustm ents (Cheatham, 2000: 587). Nothing on this topic was mentioned by Dixon.

According to a study done by Looney *et al.*, the medical schools in the United States that have a Biostatistics course cover the following topics: only 41% of the schools cover ANOVA and 31% cover multiple comparisons (Looney *et al.*, 1998:93).

Only one university indicated that they teach the concept of meta-analysis. I suggest that the other universities must try to incorporate the topic in their curricula.

4.2.13 Correlation and regression

The topics of scatter diagrams and "confidence interval and / or significance test for the correlation coefficient" are taught at 4 universities in South Africa. Fewer universities covered the rest of the topics (method of least squares, regression, multiple regression, logistic regression and comparisons of assumptions between correlation).

Confidence intervals and the fit to regression line are taught at the University of Dundee, Scotland (Sprent, 2003:522). One of the video topics at the University of Sheffield in the United Kingdom is on bivariate data. This topic covers regression and correlation. This is given to the first year students (Campbell, 2002:4). Confidence interval teaching is also a topic that is covered by a medical statistics course at the University of Sydney, Australia (Simpson, 1995:202). Regression to the mean is a topic that is covered by the medical statistics course at the University of Sydney, Australia (Simpson, 1995:202). The topics of correlation and regression are introduced to the students at the University of Tehran, Iran (Bazargan, 2002:3).

According to Dixon the following are critical under correlation and regression:

- indicate when confidence intervals have advantages over significant test (confidence intervals rather than p-values)

- explain the meaning and purpose of correlation and of least squares linear regression (scatter diagrams, correlation, regression)
- test a correlation coefficient for statistical significance (the p-value, significance levels)

and these form part of the core curriculum used at the University of Sheffield, United Kingdom (Dixon, 1994:64).

According to a study done by Looney *et al.*, the medical schools in the United States that have a Biostatistics course cover the following topics: 85% of the schools cover correlation and 69% cover linear regression (Looney *et al.*, 1998:93).

The topics under “correlation and regression” are very important according to the literature. Four of the universities in South Africa teach scatter diagrams and correlation and three universities introduce regression, so for the others that do not teach it to the students they need to take steps to incorporate these topics.

4.2.14 Other topics

Seven universities in South Africa cover the topic of critical reading. Fewer universities covered the rest of the topics (choosing the statistical method and quasi-experimental designs).

One of the most important topics of any statistical teaching found in the literature is critical reading. According to Barnett *et al.* at the Mount Sinai School of Medicine, New York, the preclinical years introduce the students to medical informatics and the skills needed to appraise the literature (Barnett *et al.*, 2000: 164).

It is important for the students to read literature and to reinforce to the students the importance of reading journal articles (Simpson, 1995: 202).

According to Hazlett, the use of information technologies (to search medical literature databases) and critical appraisal skills (to evaluate and select retrieved articles) will enable a clinician to find the manageable amount of valid information (evidence) that is relevant to one's practice (Hazlett, 1998:184).

It is clear that the importance of critical reading is seen in South African, as all participating universities introduced this topic and at most universities it had to be used.

4.2.15 Additional topics

According to Barnett *et al.* at the Mount Sinai School of Medicine, New York, the teaching of online search strategies a very important technique that students need to acquire to make their searches quicker and more efficient especially in line with the learning of EBM (Barnett *et al.*, 2000: 164). As was indicated in the "critical reading" topic, Hazlett identifies the need that online-searches has on better critical appraisal (Hazlett, 1998: 184). At the Zagreb University School of Medicine, Croatia, the topic of accessing medical literature and bibliographic databases (online searches) is taught to the medical students (Marusic and Marusic, 2003: 1236). Only one university indicated that they do data searches, but maybe one can assume that three universities are doing it because at three universities students have to do a research project. This is a topic that I would recommend that universities add to their learning programme. It is possible that students are already taught these skills during other modules (such as General Skills at the University of the Free State).

Evidence based medicine (EBM) is a critical part of the medical curriculum at the Mount Sinai School of Medicine, New York according to Barnett *et al*. (2000: 165). EBM is also considered essential at the Nuffield Department of Clinical Medicine, Oxford (Rosenberg and Donald, 1995: 1123). Only one university indicated that they do EBM. This is a topic that I would recommend to the universities to add to their learning programme. It can be that this study did not ask the right questions in regards with EBM, and thus did not collect all the necessary information.

Bayes' Theorem is also a topic that is covered by a medical statistics course at the University of Sydney, Australia (Simpson, 1995:202). According to Stangl, the only way in which we can assure that our healthcare professionals becomes better consumers of statistical information and better decision makers, is by teaching them Bayes' Theorem in the context of diagnostic testing (Stangl, 2002:1). None of the universities indicated that they introduce this topic, and this is probably a topic that is not suited for a medical statistics course.

At the Zagreb University School of Medicine, Croatia, the topic of report writing is taught to the medical students (Marusic and Marusic, 2003: 1236). Report writing skills are taught at two South African universities, and this is a skill that I would recommend the universities that do not teach this must add to their learning programme.

At the School of Medicine, University of California, San Diego the likelihood ratio of a positive and negative test are taught to the students (Deutsch, 2002:3). This topic is taught at one university in South Africa.

4.3 Questionnaire

The persons responsible for the teaching of the statistics / research methodology courses are a doctor (2 universities), statistician (6 universities) and Applied Mathematics lecturer (1 university). Three of the universities incorporate the help of tutors as well. According to the literature it is critical that more doctors and tutors must be incorporated in the courses, so that the more practical side of medical statistics can be implemented (Clayden, 1990:1031).

The specific statistics or research methodology courses show a vast variety of implementation at the different universities. Two universities changed their course as little as 2 years ago (2003), two have it since the 1970's/80's and two universities since 2000. A concern is that the universities that have not changed their course since 1970's/80's would not have implemented any EBM methods in their course, and this needs to be rectified.

The most relevant statistical concepts must be taught as early as possible in the course. And that these skills must be used and reinforced throughout the course, otherwise they may just be learned for the first year exam and then forgotten (Astin *et al.*, 2002:1004). In South Africa only one university reinforced the course during the 3rd and 5th year, after it had been taught during the 1st year. For the other, 4 universities teach the course in the 1st year, 1 in the 2nd year and 1 in the 3rd or 4th year, depending on when it is selected as an elective programme.

The stages of implementation of statistics in other modules are as follows:

- In theory at the University of Sheffield the course will be reinforced in later years with project work (Campbell, 2002:4).
- Later in the medical curriculum of the University of Bristol, students will use statistical methods in Special Study Modules (Astin *et al.*, 2002:1003).

In South Africa the respondent of one university did not know whether the statistics content is implemented again. At two universities it is not implemented again and at three universities it is implemented during academic afternoons or in the special study modules or when they have classes with doctors. This aspect must get attention at the other universities.

In 1974/5 the average classic statistical course was between 20 and 30 hours long (Wakeford, 1980: 78). The University of Sheffield, United Kingdom course currently (2001/2002) consists of one introductory lecture to all the students, and five 2 hour workshops. These five workshops are small group sessions (up to 32) students, three are based on videos produced by the University of Sheffield. These videos are approximately 20 minutes each, split into four five sessions at the end of each session the students have a practical exercise to do (Campbell, 2002:4). The University of Bristol, United Kingdom course consists out of 13 sessions that cover both medical statistics and epidemiology. Each session consists out of a small group tutorial based around one or more problems, which contain a mixture of statistical and epidemiological material (Astin *et al*., 2002:1003). In South Africa the contact sessions varies from 40 minutes to 180 minutes, and from 5 to 20 contact sessions.

The class size at the University of Sheffield is around 250 students (Campbell, 2002:4). According to Bazargan, 24 would be the ideal number of participants for a workshop on medical education research methodology (Bazargan, 2002: 2). In South Africa the class sizes are very similar for 3 universities where the average size is around 200. One university has around 140 students; one around 320 students, and one with 150 students (but the attendance will be around 40 students because this is part of the elective programme).

Students at the Royal Free & University College Medical School, United Kingdom attend practical classes. These practical classes are done in a ratio of two tutors to 25 students (Morris, 2002: 974). Four universities in South Africa use practical classes and 3 universities use tutors. I will suggest that the other universities must also try to implement practical classes as well as the use of tutors.

At the Zagreb University School of Medicine, Zagreb, Croatia the students need to work on research projects (Marusic and Marusic, 2003: 1236). At Stanford University School of Medicine, California, students are encouraged to conduct research. At completion of their research experience, the majority of students felt they had learned to formulate research questions, analyse data, use a new technique, develop methodology, review the literature critically, write a manuscript, and incorporate the guidelines for responsible conduct of research into their work (Jacobs and Cross, 1995: 342). Three universities in South Africa use research projects during their medical education. I will recommend that all the universities try to incorporate research projects, because this will show the students how important statistics is in the course and they will learn how to use statistics and research principles in the medical area.

The problems that are encountered in the presentation of the statistical / research course according to the literature are the following:

- The aim of the course is strictly limited, because it covers only a few topics in 5 work sessions of 30 minutes (Campbell, 2002:4).
- It is difficult to maintain quality control and consistency with a variety of tutors (Campbell, 2002:4).
- Do not have enough qualified tutors to cover small group work (Campbell, 2002:4).
- Limited supply of qualified staff and suitable teaching rooms (Astin *et al*., 2002:1003).

These problems seem quite different from the problems mentioned in South Africa (Table 3.18). The four problems areas are (1) students, (2) data, (3) presentation and (4) time.

Up-to-date data from current research need to be used as examples so that relevant and interesting problems are presented to the students. This could be obtained from recent journal articles or from researchers.

Some presentations problems where highlighted, and this can be an indication that the respondents (lecturers) need to rethink ways to make the statistical course more applicable and likeable to the medical students.

It seems that time is an important issue and with the shortening of courses, time will become less. This may be an area where time can be saved by cutting out unnecessary topics. The danger also exists that the teaching of statistics and research methodology is made less or even scrapped if time becomes too limited.

According to a study done by Looney *et al*. (1998:93), only 14 of the 74 medical schools that cover a Biostatistics course, covered instructions how to use computers. Only 7 of the 14 medical schools covered statistical software, 6 medical schools covered database management software, 4 medical schools covered spreadsheets, and 3 covered computerized library searches. In South Africa five of the universities expose the students to computer programs, directly in practical classes and indirectly through the research projects that the students must do. Three of the universities do not give exposure to any statistical software package and one university does not give exposure to any computer program (not Excel or any statistical software package). I would recommend that the medical students must be exposed to computer programs, like Excel, and if not in the statistics course this must happen elsewhere in the curricula.

Again it is interesting to see the variation between the student views on the research methodology / statistics learning programme in the literature versus in South Africa, the latter as reported by the respondents (Table 3.20). Two views in the literature are:

- Tutorial sessions are based on real problems, and this example driven approach quickly persuaded many students that the state-of-the-art in clinical medicine was not fixed (Astin *et al*., 2002:1003).
- The course is seen as a positive from the students and they show a positive attitude towards scientific research and EBM (Marusic and Marusic, 2003: 1236).

Some suggestions by the respondents on improvements of the course are listed in Table 3.23. Three of the suggestions surround timing. Timing of the course in terms of when during the medical curricula it must be taught, timing in the sense of time allocation to the statistical component of the curricula and timing in the sense of running programs together. The other suggestions are easier to implement like more practical examples and more critical reading, and here the respondents or lecturers can take initiatives and start incorporating it. The suggestion of Web CT is a good idea, although costly, and is already being implemented at another university with success.

4.4 Learning Programme Material

In section 3.4 the learning programme material was discussed. It gave the aims and objectives of five South African universities with regards to their statistics / research methodology courses. The aims / objectives of some of the universities according to the literature are:

- that students understand concepts and the reporting and interpretation of statistical methods (Astin *et al*., 2002:1003).
- To convey that study design and analysis techniques in clinical research studies can greatly influence the results (Deutsch, 2002: 2).
- To illustrate the process of forming a statistical analysis plan when conducting medical research (Deutsch, 2002:2).

The aims and objectives of the South African universities seem on par with what is proposed in the literature. Much attention is given to technicalities such as calculation of means and medians, the question arises whether the aims and objectives are reached fully.

4.5 Shortcomings of the study

The checklist and questionnaire had certain shortcomings.

No specific question was asked on the introduction or teaching of evidence-based medicine. Therefore the study does not give an indication to what extent evidence based medicine was part of the statistics curriculum or part of the medical curriculum at any of the medical schools in South Africa. Similarly, not enough information was collected regarding critical appraisal and how broadly it was covered at the medical schools.

More structured questions should have been asked on different modes of teaching statistics and research methods, with a focus on problem based learning (which is now mandated in undergraduate medical curricula).

Rather than relying on respondents reporting their opinion on student views of the courses, I should have enquired whether they ask students to formally assess the courses, and if so, have reported such assessment results.

An extra question that I should have included in the questionnaire is the following: Advantages of the course you are teaching?

The heading "need to calculate and use concepts" in the checklist was not appropriate to all the topics, and could have been changed to "need to use in own research" in some cases.

After the universities indicated that they teach some of the extra topics, I should have contacted all the universities again to ask whether the rest of them teach the other topics as well.

It would be of value to have more up to date information of universities around the world, regarding their teaching of medical statistics and research methodology. However, searches on Pubmed, Google and the library websites of the University of Free State, University of Stellenbosch, University of Pretoria and University of South Africa did not provide more recent published information than that reported on in this thesis.

Chapter 5

Conclusion and Recommendations

5.1 Conclusion

Five universities in South Africa give a formal Biostatistics / Research Methodology course, one university teaches a biostatistics course as part of an elective programme and one university teaches a few topics of Statistics / Research as part of a research block. One university did not participate in the study.

In South Africa the specific statistics or research methodology courses show a vast variety of implementation dates at the different universities. Only one university reinforced the course during the 3rd and 5th year, after it had been taught during the 1st year. For the other, 4 universities teach the course in the 1st year, 1 in the 2nd year and 1 in the 3rd or 4th year, depending on when it is selected as an elective programme. The class sizes vary from 40 to 320 students. Four universities use practical classes and 3 universities use tutors. Three universities use research projects during their medical education. Five of the universities expose the students to Excel, directly in practical classes and indirectly through the research projects that the students must do. The aims and objectives of the South African universities seem on par with what is proposed in the literature.

The persons responsible for the teaching of the statistics / research methodology courses are a doctor (2 universities), statistician (6 universities) and Applied Mathematics lecturer (1 university).

The following topics are taught to the medical students at most universities in South Africa:

(1) Study designs in medical research.

(2) Exploring and presenting data.

(3) Summarising data.
(4) Probability.
(5) Sampling.
(6) Statistical inference.
(7) Analysis of cross tabulation.
(8) Critical reading.

Four universities teach the topic “From sample to population”, “Analysis of the means of small samples”, scatter diagrams and correlations. Only three universities teach the topic of regression.

Survival analysis and multiple comparisons are not seen as a core topic in the medical curricula.

5.2 Recommendations

Only the fundamentals need to be introduced to the medical students, and there must not be too much extra course work added without removing unnecessary course work. The topics listed below should be introduced and one should move away from exactly how all the statistical functions and formulae work and are calculated.

The following recommendation are made:

- The statistics and research methodology teaching is integrated with other modules.
- More doctors must be involved with the teaching of medical statistics and research.
- All the types of study designs need to be introduced, because the subject is such an important one.
- The topics of incidence and prevalence should be part of the medical curricula.
- More focus be placed on the different types of sampling techniques.
- The topics of sensitivity and specificity are introduced.

- Standard error in general should be discussed.
- Analysis of the means of small samples is introduced.
- Analysis of cross tabulation is introduced.
- Introduction to meta-analysis to all universities.
- Regression, correlation and scatter diagrams are introduced.
- EBM needs to be implemented, if it has not been implemented yet.
- Report writing skills must be taught to all medical students.
- All students need to be exposed to research projects.

Further studies in the field should cover the following:

- Opinions of doctors and medical students regarding the statistics and research methodology training.
- Determining whether medical students and doctors understand and can interpret statistical analysis reported in the literature.
- Determining the methods used to assess the knowledge of medical students regarding statistics and research methodology.
- Statistics and research methodology training of other health professionals such as optometrists, dentists and nurses.

References

Altman, D.G. (1991). Practical Statistics for Medical Research, London: Chapman and Hall.

Appleton, D.R. (1990). What statistics we teach medical undergraduates and graduates? Statistics in Medicine, 9, 1013 – 1021.

Astin, J. *et al.* (2002). Medical students' perspective on the teaching of medical statistics in the undergraduate medical curriculum. Statistics in Medicine, 21, 100 –1006.

Barnett, S.H. *et al*. (2000). An integrated program for evidence-based medicine in medical school. The Mount Sinai Journal of Medicine, 67(2), 163 – 168.

Bazargan, A. (2002). Teaching statistics to medical doctors through research methods: A case of medical education research in Iran. The sixth International Conference on Teaching Statistics, Durban, South Africa.

Burge, S.M. (2003). Undergraduate medical curricula: are students being trained to meet future service needs? Clinical Medicine, 3, 243 - 246.

Campbell, M.J. (2002). Statistical training for doctors in the UK. The sixth International Conference on Teaching Statistics, Durban, South Africa.

Cheatham, M.L. (2000). A structured Curriculum for Improved Resident Education in Statistics. The American Surgeon, 66, 585 – 588.

Clayden, A.D. (1990). Who should teach medical statistics, when, how and where should it be taught? Statistics in Medicine, 9, 1031 – 1037.

Dawson-Saunders, B. and Trapp, R.G. (1994). Basic & Clinical Biostatistics, 2nd ed., United States of America: Prentice-Hall International.

Deutsch, R. (2002). Usage of medical journal articles in training for residents. The sixth International Conference on Teaching Statistics, Durban, South Africa.

Dixon, R.A. (1994). Medical statistics: content and objectives of a core course for medical students. Medical Education, 28, 59 - 67.

Evidence-Based Medicine Working Group (1992). Evidence-Based Medicine: A New Approach to teaching the practice of medicine. The Journal of the American Medical Association, 268, 2420-2425.

Fritsche, L. *et al*. (2002). Do short courses in evidence-based medicine improve knowledge and skills? Validation of Berlin questionnaire and before and after study of courses in evidence based medicine. British Medical Journal, 325,1338-1341.

Garcia-Berthou, E. and Alcaraz, C. (2004). Incongruence between test statistics and P-values in medical papers. BMC Medical Research Methodology, 4(13), 1-5.

General Medical Council (1993). Tomorrow's Doctors: Recommendations on undergraduate medical education, London.

General Medical Council (2003). Tomorrow's Doctors: Recommendations on undergraduate medical education, London.

Hazlett, C.B. (1998). Teaching evidence-based medicine. Hong Kong Medical Journal, 4, 183-190.

Health Professions Council of South Africa. (September 1999). Education and Training of Doctors in South Africa: Undergraduate Medical Education and Training, 1-20.

Hunponu-Wusu, O.O. (1977). The need for medical statistics in the training of health personnel. Medical Education, 11(5), 351-354.

Ito, P.K. (2001). Introductory Statistics Courses for non-statistics majors at Colleges and Universities. International statistical institute, 53rd session, Seoul, Korea.

Jacobs, C.D. and Cross, P.C. (1995). The value of medical student research: the experience at Stanford University School of Medicine. Medical Education, 29, 342-346.

Joubert, G. *et al.* (1999). How to write a protocol – A manual for beginner researchers. Department of Biostatistics, University of Free State.

Lejarraga, H. *et al*. (1998). A countrywide programme of continuing professional development in Argentina. Archives of Disease in Childhood, 78, 562-566.

Looney, S.W. *et al*. (1998). An update on Biostatistics Requirements in U.S. Medical Schools. Academic Medicine, 23, 92-95.

Marusic, A. and Marusic, M. (2003). Teaching students how to read and write science: a mandatory course on scientific research and communication in medicine. Academic Medicine, 78(12), 1235-9.

McKimm, J. and Jollie, C. (2004). Medical Education and Training in South Africa – Extract from international overview of medical education and training prepared as part of national Preparation for Practice project led by Universidade Nova de Lisboa, Faculdade de Ciências Medicas, Portugal, 1-18.

McLucas, K. (2003). A student's guide to medical literature (http://denison.uchsc.edu/SG/main.html), 12/15/2003.

Morris, R.W. (2002). Does EBM offer the best opportunity yet for teaching medical statistics? Statistics in Medicine, 21, 969 – 977.

Murray, G.D. (1990). How we should approach the future. Statistics in Medicine, 9, 1063 – 1068.

Palmer, C.R. (2002). Discussion: Teaching hypothesis tests: time for significant change? Statistics in Medicine, 21, 995 – 999.

Republic of South Africa (1998). Government Gazette, 18787, Chapter 2, 7(3).

Rosenberg, W. and Donald, A. (1995). Evidence-based medicine: an approach to clinical problem solving. British Medical Journal, 310, 1122-1126.

Simpson, J.M. (1995). Teaching Statistics to Non-specialists. Statistics in Medicine, 14, 199 – 208.

Sprent, P. (2003). Statistics in medical research. Swiss Medical Weekly, 133, 522-529.

Stander, I. (1999). Teaching conceptual versus theoretical statistics to medical students. International Statistical Institute, 52nd session, Helsinki, Finland.

Stangl, D. (2002). From testing to decision-making: Changing how we teach statistics to health care professionals. The sixth International Conference on Teaching Statistics, Durban, South Africa.

Sterne, J.A.C. (2002). Teaching hypothesis tests – time for significant change? Statistics in medicine, 21, 985-994.

Straus, S.E. and McAlister, F.A. (2000). Evidence-based medicine: a commentary on common criticisms. Canadian Medical Association Journal, 163(7), 837-840.

Wakeford, R.E. (1980). Report: Association for the Study of Medical Education – Statistics in Medicine. Medic al Education, 14, 73 – 79.

Watson, J.M. (2002). Doing research in statistics education: More than just data. The sixth International Conference on Teaching Statistics, Durban, South Africa.

http://www.sama.co.za/page.asp?pageid=9, SAMA CPD Unit,
Professional Development, 12/05/2004.

Appendix A: Checklist

Please indicate to which extent the following topics are taught in your statistics / research methodology learning programme or contact sessions.	Introduce to concept		Need to calculate and use concept		Need to interpret the concept		When is this taught?	Who teaches this?
TOPICS	Yes	No	Yes	No	Yes	No		
1. Introduction to Medical research								
2. Study Designs in medical Research								
2.1 Observational studies								
2.1.1 Cross sectional surveys 2.1.2 Cohort studies								
2.1.3 Case control studies								
2.1.4 Case series								
2.2 Experimental studies								
2.2.1 Controlled trials								
2.2.1.1 Randomised controlled clinical trials								
2.2.1.2 Not randomised								
2.2.1.3 Sequential trials / or interim analysis								
2.2.1.4 External controls 2.2.2 Uncontrolled trials								

TOPICS	Introduce to concept		Need to calculate and use concept		Need to interpret the concept		When is this taught?	Who teaches this?
	Yes	No	Yes	No	Yes	No		
2.3 Association and causality								
2.4 Retrospective vs. prospective 2.5 Longitudinal vs. prevalence studies								
3. Exploring & Presenting Data								
3.1 Scales of Measurement								
3.1.1 Nominal scales								
3.1.1.1 Dichotomous / binary								
3.1.1.2 Qualitative / categorical observations								
3.1.1.3 Percentages / proportions								
3.1.2 Ordinal scales								
3.1.3 Numerical scales								
3.1.3.1 Quantitative								
3.1.3.2 Continuous 3.1.3.3 Discrete								
3.1.4 Tables and graphs for nominal and ordinal data								
3.1.4.1 Frequency distribution								
3.1.4.2 Bar charts								
3.1.4.3 Contingency table								
3.1.5 Tables and graphs for numerical data								
3.1.5.1 Stem and leaf plots								

TOPICS	Introduce to concept		Need to calculate and use concept		Need to interpret the concept		When is this taught?	Who teaches this?
	Yes	No	Yes	No	Yes	No		
3.1.5.2 Frequency tables								
3.1.5.3 Histograms 3.1.5.4 Box and whisker plots								
3.1.5.5 Pie charts								
4. Summarising Data								
4.1 Measures of the middle (Central tendency)								
4.1.1 Mean								
4.1.2 Weighted average								
4.1.3 Median								
4.1.4 Mode								
4.1.5 Geometric mean								
4.2 Measures of spread								
4.2.1 Range 4.2.2 Standard deviation								
4.2.3 Variance								
4.2.4 Coefficient of variation								
4.2.5 Standard error								
4.2.6 Percentiles								
4.2.7 Interquartile range								
4.3 Measures to use with nominal data								

TOPICS	Introduce to concept		Need to calculate and use concept		Need to interpret the concept		When is this taught?	Who teaches this?
	Yes	No	Yes	No	Yes	No		
4.3.1 Proportions and percentages								
4.3.2 Ratios and rates 4.3.3 Vital statistics rates								
4.3.3.1 Mortality rates								
4.3.3.2 Morbidity rates								
4.3.3.3 Prevalence								
4.3.3.4 Incidence								
4.3.3.5 Adjusted rates								
4.3.3.5.1 Direct method of adjusting rates								
4.3.3.5.2 Indirect method of adjusting rates								
4.4 Measures to describe relationships between two characteristics								
4.4.1 Describing the relationship between two characteristics								
4.4.1.1 Correlation coefficient								
4.4.1.2 Coefficient of determination 4.4.2 Describing the relationship between two ordinal characteristics								
4.4.2.1 Spearman rank correlation								
4.4.3 Describing the relationship between two nominal characteristics								
4.4.3.1 Relative risk								
4.4.3.2 Odds ratio								
4.5 Reliability								

TOPICS	Introduce to concept		Need to calculate and use concept		Need to interpret the concept		When is this taught?	Who teaches this?
	Yes	No	Yes	No	Yes	No		
5. Probability								
5.1 Random variables & probability distributions								
5.1.1 Binomial distribution								
5.1.2 Poisson distribution								
5.1.3 Normal distribution								
6. Sampling								
6.1. Sampling								
6.1.1 Random sampling								
6.1.2 Systematic sampling								
6.1.3 Stratified sampling								
6.1.4 Cluster sampling								
6.1.5 Non-probability sampling								
7. Clinical Measurement								
7.1 Repeatability and precision in measurements								
7.2 Digit preference								
7.3 Comparing two methods of measurement								
7.4 Sensitivity and specificy								
7.5 Positive predictive value								
7.6 Negative predictive value								

TOPICS	Introduce to concept		Need to calculate and use concept		Need to interpret the concept		When is this taught?	Who teaches this?
	Yes	No	Yes	No	Yes	No		
7.7 Accuracy								
7.8 Receiver operating curves (ROC)								
7.9 Survival data								
7.10 Computer aided diagnosis								
7.11 Kaplan-Meier curves								
7.12 Life-table analysis								
7.13 Actuarial analysis								
8. From Sample to Population								
8.1 Properties of the normal distribution								
8.2 Reference ranges								
8.3 Standard error of a sample mean								
8.4 Standard error of a proportion								
8.5 Standard error of the difference between two means								
8.6 Standard error of the difference between two proportions								
8.7 Sample size for an estimate								
9. Statistical inference								
9.1 Testing a hypothesis								
9.2 The null hypothesis								
9.3 Type one/two errors								

TOPICS	Introduce to concept		Need to calculate and use concept		Need to interpret the concept		When is this taught?	Who teaches this?
	Yes	No	Yes	No	Yes	No		
9.4 The p-value								
9.5 Significance levels								
9.6 One-and two-sided tests of significance								
9.7 Significant, real and important								
9.8 Statistical power								
9.9 Degrees of freedom								
9.10 Confidence intervals rather than p-values								
9.11 Comparing the means of large samples using the normal distribution								
9.12 Comparison of two proportions								
9.13 Multiple significance tests								
9.14 Exploratory data analysis								
10. Analysis of the means of small samples								
10.1 The t distribution								
10.2 The one sample t method								
10.3 The means of two independent samples								
11. The analysis of cross tabulations								
11.1 The chi-squared test for association								
11.2 Validity of the chi-squared test for small samples								

TOPICS	Introduce to concept		Need to calculate and use concept		Need to interpret the concept		When is this taught?	Who teaches this?
	Yes	No	Yes	No	Yes	No		
11.3 Fisher's exact test								
11.4 Yates` continuity correction for a 2 by 2 table								
11.5 McNemar`s test for matched samples								
12. Methods based on rank order								
12.1 The Mann-Whitney U test								
12.2 Wilcoxon matched pairs test								
12.3 Kendall's rank correlation coefficient								
13. Multiple comparisons								
13.1 ANOVA								
13.2 Bonferroni adjustments								
13.3 Meta-analysis								
13.4 Post Hoc comparisons								
14. Correlation and regression								
a. Scatter diagrams								
b. Confidence interval and / or significance test for the correlation coefficient								
c. The method of least squares								

TOPICS	Introduce to concept		Need to calculate and use concept		Need to interpret the concept		When is this taught?	Who teaches this?
	Yes	No	Yes	No	Yes	No		
d. Regression								
e. Comparison of assumptions between correlation and regression								
f. Multiple regression								
g. Logistic regression								
15. Other topics								
a. Choosing the statistical method								
b. Critical reading								
c. Quasi - experimental designs								
d. Quantitative designs								
16. Any additional topics								
Sources used in the compilation of the checklist:								

.==> Altman, D.G. (1991). Practical Statistics for Medical Research.

.==> Cheatham, M.L. (2000). A structured curriculum for improved resident education in Statistics. The Americ an Surgeon, 66, 585-588

.==> Dawson-Saunders, B. (1994). Basic & Clinical Biostatistics.

.==> Dixon, R.A. (1994). Medical statistics: content and objectives of a core course for medical students. Medical Education, 28, 59 - 67.

Appendix B – Questionnaire

1. When during the curriculum are the statistics / research methodology learning programme (contact session) taught to medical students? (Example first semester first year.)

……………………………………………………………………………………

……………………………………………………………………………………

2. Who teaches this statistics and / or research methodology-learning programme (contact session)?
   a. Theoretical Statistician
   b. Applied Statistician
   c. Medical Profession
   d. Other (please indicate)…………………………………………………………………..

3. a. How long are the contact sessions?

   ………………………………………………………………………………

   b. How many contact sessions do you have?

   ………………………………………………………………………………

   c. How big are your classes (Number of students)?

   ………………………………………………………………………………

4. Does groupwork take place?
   a. Yes / No
   b. If yes, briefly discuss the process.

   ………………………………………………………………………………

   ………………………………………………………………………………

   ………………………………………………………………………………

   ………………………………………………………………………………

   ………………………………………………………………………………

5. Are there practical classes?

   a. Yes / No

   b. If yes, briefly discuss the process.

   …………………………………………………………………………………

   …………………………………………………………………………………

   …………………………………………………………………………………

   …………………………………………………………………………………

   …………………………………………………………………………………

6. Do you use the help of tutors?

   a. Yes / No

   b. If yes, who act as tutors?

   …………………………………………………………………………………

   …………………………………………………………………………………

   …………………………………………………………………………………

7. What problems do you encounter in,

   a. The course itself?

   …………………………………………………………………………………

   …………………………………………………………………………………

   …………………………………………………………………………………

   …………………………………………………………………………………

   …………………………………………………………………………………

   b. The presentation of the research methodology / statistics learning programme (contact session)?

   …………………………………………………………………………………

   …………………………………………………………………………………

   …………………………………………………………………………………

   …………………………………………………………………………………

   …………………………………………………………………………………

8. Must the students do a research project?

a. Yes / No

b. If yes, briefly discuss the process.

………………………………………………………………………………………

………………………………………………………………………………………

………………………………………………………………………………………

………………………………………………………………………………………

………………………………………………………………………………………

9. Are the students exposed to computer work?

a. If yes, what computer programs?

i. SAS

ii. Statistica

iii. SPSS

iv. S-Plus

v. NCSS

vi. NVino

vii. Other (Please name them)

……………………………………………………………………….

10. In your experience how do students view this module?

………………………………………………………………………………………

………………………………………………………………………………………

………………………………………………………………………………………

………………………………………………………………………………………

11. What teaching methods do you utilize in presenting the module?

………………………………………………………………………………………

………………………………………………………………………………………

………………………………………………………………………………………

………………………………………………………………………………………

12. Are the statistics and research methodology taught as a separate module or an integrated module?

……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………

13. Is the theory applied in other modules, and at what stage?

……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
…………………………………………………………………………
……………………………………………………………………………………………
…………………

14. What can be improved in the course?

……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………

15. Do you have any other suggestions?

……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………

Appendix C – Consent form (Afrikaans) - Heads of Schools

Inligtingsbrief

Ek is ‘n M.Med.Sc. student (2004154353) geregistreer by die departement van Biostatistiek by die Universiteit van die Vrystaat en saam met professor Gina Joubert (my studie leier) is ons besig om ‘n studie oor die profiel van statistiek en navorsingsmetodiek van voorgraadse mediese studente by Suid-Afrikaanse universiteite te doen. Ons wil my bevindinge van die Suid-Afrikaanse universiteite vergelyk met wat in die literatuur vervat word. Die waarde van die studie is dat universiteite in Suid-Afrika met mekaar vergelykbaar is. Voorstelle sal gemaak word oor die verbetering van kontaksessies en “learning programmes”. Die studie sal een einde 2005 voltooid wees.

Toestemmingsbrief

Met die volgende bevestig ek dat ek volledig ingelig is oor die beoogde studie en dat my deelname heeltemal vrywillig is. Ek weet dat alle informasie as vertroulik hanteer sal word.

Universiteit: ……………………………………

Naam: …………………………………………

Datum:…………………………………………..

Persoon(e) verantwoordelik vir die aanbied van die statistiek en / of navorsingmetodiek by die mediese skool.

Naam Tel epos

……………………………………………………………………………………………

……………………………………………………………………………………………

Appendix D – Consent form (English) - Heads of Schools

Information piece

I am a M.Med.Sc. student (2004154353) registered in the department of Biostatistics at the University of Free State and together with professor Gina Joubert (my study leader) we are conducting a study on the profile of statistics and research training of undergraduate medical students at South African universities. We want to compare my findings of South African Universities with what is proposed in the literature. The values of the study are that each university can compare itself to the rest of the universities in South Africa and that recommendations will be made to improve the learning programmes or contact sessions. The study will be finished at the end of 2005.

Consent form

With this I confirm that I was fully informed about the above study and that participation is voluntary. I know that all information will be handled confidentially.

University: ……………………………………

Name: …………………………………………

Date:…………………………………………..

Person (s) responsible for the teaching of statistics and research methodology at our school.

Name Tel email

………………………………………………………………………………………

………………………………………………………………………………………

Appendix E – Consent form - Respondents

Information piece

I am a M.Med.Sc. student (2004154353) registered in the department of Biostatistics at the University of Free State and together with professor Gina Joubert (my study leader) we are conducting a study on the profile of statistics and research training of undergraduate medical students at South African universities. We want to compare my finding of South African Universities with what is proposed in the literature. The values of the study are that each university can compare itself to the rest of the universities in South Africa and that recommendations will be made to improve the learning programmes or contact sessions. The study will be finished at the end of 2005.

Would you agree to participate, I will contact you to arrange a date and time that will suit you to interview and gather the relevant information to complete the checklist / questionnaire.

Your participation is voluntary and can at any time during the study be terminated. The sources of information will be kept confidential and feedback will be given to each respondent. We hope to publish the information obtained.

Consent form

With this I confirm that I was fully informed about the above study and that participation is voluntary. I know that I can terminate my participation at any time during the study and that all information will be handled confidentially.

Name: ………………………………………

Date:…………………………………………..

Signature:……………………………………..