TARANAKI MEDICAL FOUNDATION

2024

A Research Guide

Summer student projects

The Taranaki Medical Foundation annually awards scholarships for medical students for participation in research projects led by senior medical clinicians in Taranaki. These projects, spanning seven to eight weeks between November and February (with dates flexible depending on the universities and students), are open to eligible medical students from their second to the fourth year of study at the Universities of Auckland and Otago. Due to their level of training, students are not permitted to have direct clinical contact with patients (although they may observe with patient consent).

What is achievable during this timeframe depends on multiple factors:

- The respective experience in research of both the student and the supervisor,
- The clinical knowledge of the student in the area of research of the supervisor,
- The stage of the research project, notably if the research project has already reached ethics approval (no data collection should be started without that approval),
- The availability of resources for the research, such as access to full-text articles, access to medical records or databases and/or availability of research software,
- The availability of the supervisor to monitor the progress of the student,
- The availability of external resources or additional staff if needed for the research project.

These constraints should be considered by the supervisor before the electronic submission of the application to benefit from the scholarship.

Whatever the student's contribution to the research project is, supervisors should make sure that the work of the student is duly traced, easily transmitted to another person once the 8-week window has ended and reproducible. Otherwise, there is a high chance that the contribution of the student might not be useful, especially if the project has multiple stop-and-go steps.

Supervisors should be conscious that, usually, the first 1 to 2 weeks of work should be dedicated for the student to be familiarized with the clinical context (ward/ED/consultation rooms), the research team and the different tools they will be working in or with. Similarly, the last week should be dedicated to summarizing the work done by the student, offering feedback on their work and correcting any mistakes they might have potentially made before they head back to their university grounds.

We would like to remind all the supervisors that, according to good practice rules in medical publications, the work of the medical students should be recognised as co-authors for all the publications born out of the research project they were involved in. The place in the rank of authors should depend on their overall contribution.

This guide aims at providing some examples of what is achievable by a medical student with little to no research experience, working full-time, during the 8 weeks of a typical scholarship.

Designing a project

The student can help the supervisor write down a comprehensive research project so that it can be published in a trials database/journal, submitted for approval by an ethics committee or used to get research grants.

The first step would be for the student to organise a literature review of what has already been published on the topic and write down a summary of their findings to be recontextualised by the supervisor. This supposes that the student knows or learns how to use publication databases (such as MEDLINE), collaborates with the hospital librarian if possible and produce a first database of relevant articles. Depending on the breadth of the research project, this work can take from 2 to 4 weeks. If the research project is actually a publishable literature review, it might take the whole 8 weeks to do a first collection of relevant articles and apply the selection criteria for these articles.

The work of the student can also consist of determining what is the best methodology for the research question. Depending on their statistical skills, they can liaise with statisticians and methodologists to determine how many subjects to recruit, how to recruit them effectively, how to randomize (if necessary), work out cluster sizes (if it is a multicentric study for example), determining the duration of the inclusion process, etc. Depending on the size of the study, this work can take from to 2 to 4 weeks.

At the design phase, the student's work can also be about the tools of the research project. For example, a student can help create a relevant case report form (CRF) for a prospective study and test it on a small number of subjects. Or they can design a questionnaire that would be given to the patients after consultation or operation, test it on a small sample, and refine it. Another possibility would be for them to create the data collection template (in Excel or in statistical software), which would be used later in the collection phase. They can also help build content for a website related to the study. This type of work could typically take 4 to 8 weeks.

Finally, the student can concretely type the research project for later submission based on the different elements listed above. This would need tight supervision from the senior clinician, notably in terms of time management skills as nobody wants it to be rushed at the last minute for a deadline. Depending on the size of the project, that can take anywhere from 1 week to 8 weeks (in the case of a multicentric randomized

controlled trial). If the trial design is to be published in a journal before data collection and the student did the first draft of that publication, good ethics in publishing would recommend the student to be the first author of that trial publication (and the supervisor to be either the second or last author and the one to contact).

Collecting data

Data collection should start only once the research project has been approved by an ethics committee. Depending on the design of the study, there are multiple ways data can be collected, and medical students can either participate individually in this data collection or coordinate it.

As a first example, the design of the study can lead to collecting information in past medical records. The student's task can then be requesting the medical records from the hospital archives, checking for quality criteria, reading the records and collecting the data in a structured way, and complementing the data by additional research (for example, extending the data collection to primary care providers records). This can also be applied to database work (for example, a joint replacement registry). The amount of work needed will vastly depend on how targeted the information search is. For example, if the only information we want to know is whether people have any record of smoking, this can be applied very easily to hundreds of patient records. On the contrary, if there is some complexity, such as trying to spot a chain of events leading to a medical complication or deciphering complex interactions between medications, it may take up to half a day per medical record. Therefore, this data collection can easily last for the whole 8 weeks of the scholarship. What is absolutely necessary is that the data is collected in a consistent and traceable manner so that, if the collection is not finished at the end of the 8 weeks, the work can be continued by another researcher. We recommend that the data be collected in Excel or database software with very well-defined criteria.

Another example would be for the medical students to collect data through questionnaires, interviews, or focus groups. In that case, a lot of time will be spent through face-to-face (or at least phone/Zoom) contact with the subjects of the study. Paper questionnaires can be handed and collected later, but considering the low rate of response of that method, the time saved by not doing a face-to-face interview is usually lost in tracking missing questionnaires or ensuring a better response rate. For qualitative research, notably the use of semi-structured interviews, a considerable amount of time will be dedicated to interviewing the

subjects of the study, and the student must ensure they have enough flexibility to arrange the interviews at a convenient time for the respondents (perhaps also in the evenings or weekends). All data collected by these methods should be compiled by the student in a database to be used at a later stage of the research. Once again, depending on the size of the study, this collection work can last up to the full 8 weeks of the project.

A third example of data collection by the medical students is the use of CRFs (case report forms) in a prospective study. While similar to the methods described in the previous paragraph, CRFs typically involve a degree of clinical examination and the collection of additional investigations (laboratory, radiology, etc.). However, because the medical students are not allowed to have direct clinical contact with patients, they would not be able to conduct any clinical examinations themselves. However, with patient permission, they may sit in on clinics and ask survey questions but are not allowed to touch any patients. The supervisor should ensure that the necessary clinical skills for the examination are mastered by qualified personnel and that the clinical findings are reproducible with different operators. Overall, using CRFs is expected to be a slower process than handing out questionnaires or doing interviews. On average, a thorough completion of CRF is expected to last around 1 to 2 hours per patient – but once again, this depends on the level of detail that is required.

A last example is the case of a medical student coordinating data collection. The student can be tasked to monitor the quality of data collection among different examiners or to make sure that all patients who should be included in the study have been contacted and offered to participate in the study. For instance, if the student is involved in an ED study that requires the ED staff to include patients and ask a few patients when they present with certain conditions, the medical student can coordinate the initial training of all the ED staff, review attendance lists every day to make sure inclusion is maximized, supervise the quality of the data collection by the different staff members and feedback to them.

Analysing data

If the student gets involved after the data collection process is completed, they can help with the analysis of the data. This usually requires the student to have previous experience in research and/or statistical skills. We do not recommend that a “naïve” student on these matters should be given any responsibility of data analysis unless they can be monitored full-time by an experienced researcher or statistician.

Often, the first step of an analysis is data cleansing, i.e. detecting and removing/correcting abnormal entries from the dataset. This process needs to be done in a very cautious way, with perfect traceability (for quality control and reproducibility purposes), and multiple saves, as once a record is deleted from a database, it is very hard to retrieve it. and it can skew the results quite sensibly. As long as the student understands the impact of this data cleansing and has received training for it, this is an achievable objective for a summer scholarship. Depending on the size of the database, it can take from 2 to 8 weeks to be achieved.

In terms of pure analysis, depending on the skills of both the students and the supervisor, tasks can range from formatting tables (notably using pivot tables functions in Excel) or working on graphical representations of the data to determining a multivariate equation to explain the findings of the study. If the aim of the student contribution is, for example, to get a nice graph to be displayed on a research poster, it should be done within 1 or 2 weeks. If the contribution is to do a full statistical analysis, the average time that would be required for that would be 3 to 4 weeks.

Writing a publication

Finally, students can be tasked with writing a first draft of a publication, whether it is a full article aimed at being published in a peer-reviewed journal, slides for a research conference, a public health report or a summary of the current state of the art of a medical procedure in Taranaki.

Writing down a full article, from introduction to conclusion, is a hard task that would keep the student busy, likely for the 8 weeks, especially if they have not been involved in the research project initially. But it is a very rewarding task that will help them all along their career for their training papers or if they want to lead a study themselves later on. A good estimation of work division would be 2 weeks for the student to familiarise themselves with the study context and findings, 4 weeks for writing the first draft, and 2 weeks for back-and-forth exchanges with their supervisor. Writing a paper is often thought of as a quick and easy task to do and often postponed until the last minute. But experience in this field shows clearly that regular work every day is a much better way to achieve a readable paper than a last-minute writing rush. The supervisor is expected to get regular updates on the writing process from the student, our advice would be to monitor that at least twice a week.

Unless the first draft produced by the student needs a full overhaul, it is expected that a student who writes a scientific article gets credited the first place in authorship. But this comes with the responsibility for the student to remain reachable past the 8 weeks of the scholarship, notably to exchange with the rest of the co-authors and rewrite sections if need be.

Resources and tools

Hospital and university libraries, including their staff can help design database research equations and compile lists of articles for the student.

Reference manager software: Zotero is recommended for medical students, considering it is free, open-source and has vast documentation. Alternatives are Endnote or Mendeley.

Access to IT systems and software: it is recommended for the supervisor to document themselves before the start of the scholarship on how to get access for the student to the electronic medical records, to the radiology and laboratory servers or to the paper archives of medical records.

Access to Microsoft Office suite (or free alternatives such as OpenOffice, LibreOffice or Google Docs), including Excel, Word and PowerPoint as well as productivity tools such as Teams, Outlook or SharePoint servers are essential. **Secure storage of medical information is paramount** and must be defined and explained to the student before the start of any data collection or analysis. While ideally, students should not have a copy of the database on their personal laptops, the reality is that most students will have to use their own laptops, especially for after-hours or outside hospital/GP office data collection. Students should be made aware of the importance of secure data storage and should only keep anonymized numerical data on their laptops. Identifiable patient data must remain on the hospital server, and students should use their laptops to log into the hospital system remotely if no computer terminals are available.

Access to database and statistical software. Ideally, R can be used for these 2 functions, is free, open-source and vastly supported on the Internet. It is very intimidating at first, but its use can be much improved by using a GUI (Graphics User Interface) such as Bluesky Statistics or R commander. Alternatives are JASP (increasingly popular and free), Jamovi (based on R but transparent for the user) or even a website such as Datatab. It is not recommended to use Excel to do advanced statistics as it does not work through the series of reproducible instructions, it works directly on the database. Complex and very expensive software such as SAS, SPSS or STATA are probably out-of-reach for medical students unless they have previous experience with it and have access to a research lab that owns licenses for these programs.

Questionnaire tools: Microsoft Office and Google Docs have their own version of online surveys that feed an Excel file or equivalent automatically, but for more advanced complex surveys (with

conditional responses and management of e-mail databases), LimeSurvey is the most powerful tool, free-to-use if installed on your own server. Other alternatives are Surveymonkey, Qualtrics or Typeform.

Qualitative research tools: nVivo is by far the most recommended one and is free for students. Alternatives are ATLAS, maxQDA