McGill

New Program/Concentration Proposal Form

(2019)

1.0 Degree Title
Please specify the two degrees for concurrent degree programs

Graduate Certificate

2.0 Administering Faculty or GPS

GPS

1.1 Major (Subject/Discipline)(30-char. max.)

Health Data Analytics

Offering Faculty & Department

Medicine and Health Sciences/Epidemiology, Biostatistics and Occupational Health

1.2 Concentration (Option)
(30 char. max.)

3.0 Effective Term of Implementation
(Ex. Sept. 2019 or 201909)
Term

202209

1.3 Complete Program Title (from boxes 1 and 5)

Graduate Certificate in Health Data Analytics

4.0 Rationale and Admission Requirements for New Program/Concentration

Rationale: Please see attached proposal.
Admissions requirements: A Bachelor's degree in a related, quantitative discipline (including but not restricted to Biology, Life Sciences, Mathematics, Physics, Psychology).
Applicants with a Cumulative Grade Point Average below 3.2 (out of 4.0), or a Grade Point Average below 3.2 in the last two years of full-time studies will not be considered for admission.

5.0 Program Information
Please check appropriate box(es)

5.1 Program Type
- ☐ Bachelor's Program
- ☐ Master's
- ☐ M.Sc.(Applied) Program
- ☐ Dual Degree/Concurrent Program
- ☐ Certificate
- ☐ Diploma
- ☒ Graduate Certificate
- ☐ Graduate Diploma
- ☐ Ph.D. Program
- ☐ Doctorate Program (Other than Ph.D.)
- ☐ Self-Funded/Private Program
- ☐ Off-Campus Program
- ☐ Professional Development Cert
- ☒ Distance Education Program
- ☐ Other (Please specify)

5.2 Category
- ☐ Faculty Program (FP)
- ☐ Major
- ☐ Joint Major
- ☐ Major Concentration (CON)
- ☐ Minor
- ☐ Minor Concentration (CON)
- ☐ Honours (HON)
- ☐ Joint Honours Component (HC)
- ☐ Internship/Co-op
- ☐ Thesis (T)
- ☐ Non-Thesis (N)
- ☐ Other

Please specify

5.3 Level
- ☐ Undergraduate
- ☐ Dentistry/Law/Medicine
- ☐ Continuing Studies (Non-Credit)
- ☐ Collegial
- ☒ Masters & Grad Dips & Certs
- ☐ Doctorate
- ☐ Post-Graduate Medicine/Dentistry
- ☐ Graduate Qualifying
- ☐

5.4 Requires Centrally-Funded Resources Yes ✓ No __

6.0 Total Credits or CEUs (if latter, indicate "CEUs" in box)

15

7.0 Consultation with
Related Units Yes ☒ No ☐
Financial Consult Yes ☒ No ☐
Attach list of consultations.

8.0 Program Description (Maximum 150 words)

The Graduate Certificate in Health Data Analytics is an online program that offers strong statistical and computational training to solve emerging problems in public health and the biomedical sciences. This training aims to provide the skills needed to clean, manage and analyze data sets, and to interpret and report findings in a reproducible, effective, and accessible manner.

9.0 List of proposed new Program/Concentration

If new concentration (option) of an existing program, a program layout (list of all courses) of existing program must be attached.

Proposed program (list courses as follows: Subj Code/Crse Num, Title, Credit Weight under the headings of: Required Courses, Complementary Courses, Elective Courses)

Graduate Certificate in Health Data Analytics (15 credits)

Required Courses (15 credits)

BIOS 640 Introduction to Health Data Science Methods (3 cr.)
BIOS 641 Health Data Analysis: Essential Theory (3 cr.)
BIOS 642 Health Data: Statistical Learning and Model Visualization (3 cr.)
BIOS 643 Theory to Practice: Analysis and Reporting (3 cr.)
BIOS 644 Health Data Analytics Capstone Project 1 (3 cr.)

Attach extra page(s) as needed

10.0 Approvals

Routing Sequence	Name	Signature	Meeting Date
Department	Robert Platt		07 Dec 2020
Curric/Acad Committee	Melissa Vollrath - FCC Chair	member: 32254193-4556-48C0-9459-0D5F5EC802FC	31 March 2021
Faculty 1		C794CE2C-FFC5-496E-BD81-062185857860 2021.03.31 16:17:06 -04'00'	
Faculty 2	Aimee Ryan- Assoc Dean, FMHS	Aimee Ryan Digitally signed by Aimee Ryan Date: 2021.04.01 16:40:58 -04'00'	31 March 2021
Faculty 3			
CGPS			
SCTP			
APC			
Senate			

Submitted by

Name	Erica Moodie
Phone	514-398-5520
Email	erica.moodie@mcgill.ca
Submission Date	11 December, 2020

To be completed by ES:

CIP Code

Certificate in Health Data Analytics Proposal – Executive Summary

McGill Biostatistics. Prepared by Erica Moodie in collaboration with Sahir Bhatnagar, Alexandra Schmidt, Gilles Paradis, and Robert Platt and financial guidance from John Kuchyna.

1. Technical Description

The emergence of health data science presents McGill with exciting opportunities to increase the impact of Biostatistics. We propose a Graduate Certificate in Health Data Analytics designed to provide solid quantitative training and essential statistical and computing skills needed to clean, manage, and analyze health science data to address important questions in public health and biomedical sciences.

2. Strategic Position of the Project

The new School of Population and Global Health at McGill offers a unique opportunity to rethink and reorganize educational offerings across the Faculty, with a focus on expanding high quality training in Biostatistics and Health Data Analysis. As a division that has long offered training both to Biostatistics students with highly mathematical backgrounds and to students in programs including Epidemiology and Public Health where technical rigour is not required, we are ideally placed to expand our teaching to reach a new audience of students who may be unable or unwilling to join a traditional on-campus program but still seek the quality and reputation that McGill has to offer. Both the inaugural director of the School, Dr Tim Evans, and the Dean of Medicine, Dr David Eidelman, fully support this proposal, both in terms of its pedagogical mandate and as an opportunity for growth within the faculty. The proposed certificates provide an opportunity to be leaders in the McGill community in terms of developing a new approach to outreach education.

3. Comparative Advantage of the Project

a) Review of Similar Initiatives in Canada

Programs in data science have been opened across Canada (Figure 1). Among the top ranked Canadian universities, including the Universities of Toronto, Waterloo, British Columbia, Simon Fraser, Queens, and Western, McGill is the only university which does not have a data science program. However, none of these existing programs have a specific focus on health data analytics, an area where our Department offers strength, credibility, and a sound international reputation. What is missing from these data science programs (and nearly all across the world with only a few exceptions[1][2] is a focus on health science research. This provides our Department with a unique opportunity to leverage the biostatistical expertise of its faculty along with its health driven research to create its own program in Health Data Science.

There exist some related initiatives at McGill. The Department of Mathematics & Statistics is investigating a Data Science/Applied Statistics MSc, and the School of Continuing Studies launched Professional Development Certificates in both Data Science & Machine Learning, and Data Analytics for Business[3]. Again, these programs do not focus on health data and we do not envision meaningful overlap.

b) Assessment of McGill's Position Relative to Peers

As one of only three Canadian universities among the top 50 on the Times Higher Education rankings, McGill is well-recognized internationally. The Division of Biostatistics houses the largest concentration of PhD level statisticians of any like department in Canada, and is internationally recognized for its strength in teaching and research. Although we, as a University, are new to the online program format, our reputation as a leading institution of higher learning bodes well for the success of these certificates.

[1] www.hsph.harvard.edu/health-data-science/program/
[2] www.manchester.ac.uk/study/masters/courses/list/10076/msc-health-data-science/
[3] www.mcgill.ca/continuingstudies/area-of-study/data-science

c) Environmental Scan and Market Study

Interest in our Biostatistics programs has increased over the past 11 years, reflecting increasing demand of (bio)statistics trained graduates in the job market, both academic and in industry. However, the number of offers made has remained stable: we have been unable to increase enrolment to meet that need with our current programs for two reasons: (1) many applicants' background is not sufficiently statistical to succeed in our existing MSc programs, and (2) insufficient supervisory capacity in terms of both funding and time.

Projected certificate enrollment is for 15 students in the first year, 30 in the second, and a steady state of 45 thereafter. We anticipate a relatively low rate of attrition (15%) given the McGill name, the short duration of the program, and its flexibility. We expect a high uptake (at least 50%) of international students based on the demographics of applicants to our existing MSc in Biostatistics.

4. Description of Activities

Target audience & Prerequisites: The primary audience for our proposed certificate are students who wish to pursue graduate studies in Biostatistics who (i) may lack the necessary mathematical preparation for our existing graduate degrees in Biostatistics (and would similarly be unprepared for a graduate degree in Mathematics & Statistics), (ii) seek a flexible schedule, or (iii) wish to receive McGill training without the requirement of being physically present in Montreal. The only prerequisite to entry in the program is that applicants must hold an undergraduate degree in a quantitative or science field, for example biology, physics, or psychology. Applicants with a Cumulative Grade Point Average below 3.2 (out of 4.0), or a Grade Point Average below 3.2 in the last two years of full-time studies will not be considered for admission. Unlike the graduate programs in Biostatistics, applicants are not required to have taken courses in Calculus, Advanced Linear Algebra, or Real Analysis.

The program will consist entirely of asynchronous, online teaching. Virtual office hours and small group meetings will be offered to maximize access according to students' time zones and schedules; students will alternatively be permitted to submit video presentations in lieu of live presentations. This asynchronous and fully online mode of delivery and the lack of any mandatory in-person attendance will serve to accommodate those students unable to be physically present at McGill or accommodate a more traditional learning schedule.

Curriculum: We propose a 10-month, 15 credit graduate certificate which blends strong statistical and computational training to solve emerging problems in public health and the biomedical sciences. This training will enable students to clean, manage and analyze data sets, and learn how to interpret and report their findings in a reproducible way. Professional communication of analyses will be emphasized throughout the certificates, and particularly the focus of BIOS 643 and 644. All courses will be made available online, i.e., there is no on-site requirement. The proposed curriculum is shown in Table 1 below, and will run from September to June, inclusive.

Table 1. Proposed 10-month Health Data Analytics Certificate I curriculum

Term	Course Name	Credits
Fall – first half	Introduction to Health Data Science Methods (BIOS 640)	3
Fall – second half	Health Data Analysis: Essential Theory (BIOS 641)	3
Winter – first half	Health Data: Statistical Learning & Model Visualization (BIOS 642)	3
Winter – second half	Theory to Practice: Analysis and Reporting (BIOS 643)	3
Summer	Health Data Analytics Capstone Project 1 (BIOS 644)	3

While there are two courses in each of Fall and Winter term, the courses will be taken sequentially over six weeks each, rather than concurrently, to maximize course-completion and maintain engagement.

5. List of Stakeholders

The primary stakeholders in this endeavour within the SPGH are the inaugural Director, Tim Evans, and the Chair of Epidemiology, Biostatistics, and Occupational Health, Robert Platt. Additional key stakeholders are the core Biostatistics faculty members. Further stakeholders within the Faculty of Medicine are Dean David Eidelman, Associate Dean Aimee Ryan, and Vice-Dean of Education Annette Majnemer. Josephine Nalbantoglu, Dean of Graduate and Postdoctoral Studies, is also a key stakeholder in this venture. **The proposed Certificate including the draft budget were approved by the Dean's Operations Committee of the Faculty of Medicine and Health Sciences on November 16, 2020.**

We have consulted with (1) the Department of Mathematics & Statistics, (2) the School of Computer Science, (3) the Quantitative Life Sciences program, (4) the Integrated Program in Neuroscience, (5) Human Genetics, (6) the Desautels School of Management, and (7) the School of Information Studies on the content of the first certificate and all its courses. All groups have approved these proposed courses (approvals attached in Minerva). While there exist many courses in (bio)statistics and data analysis on campus and so there is some overlap with existing courses the content is overlapping but not identical and none of the existing courses aim to serve the online, non-mathematical students that we believe to be the primary audience for this certificate. None of the consulted departments that offer statistics courses had any concerns or objections on the grounds of overlap. We did receive some questions as to whether students in other programs could take these courses, further underlining the fact that these proposed courses offer content and flexibility not currently available at McGill. Students outside of the certificate would indeed be permitted to take the courses; such students would not be required to meet the strict prerequisites (following courses 640 through to 644 in sequence) but rather could be admitted by permission of the instructor if equivalent courses to the prerequisites have been completed.

6. Detailed Timetable

We propose opening enrolment to the Certificate in September 2022. Accordingly, we anticipate hiring new faculty to ensure a smooth roll-out of our certificate without straining existing programs (Biostatistics MSc/PhD, Epidemiology MSc/PhD, and Public Health MSc) which all depend on Biostatistics faculty for core courses and are already feeling the strain of two recent departures.

To meet our timeline, we anticipate securing new course approval for all five of the Certificate courses by Fall 2021. We will work with Teaching and Learning Services throughout Fall 2021 and Winter 2022 to produce the online course content for the Graduate Certificate.

7. Risks

There are two foreseeable risks to the success of the program. The first risk is that we are unable to hire new Faculty on schedule. This would create significant strain on our existing Faculty, who already carry a high teaching load while maintaining active research and supervisory programs. This strain could be alleviated by hiring short-term contract lecturers. The second risk to the success of the program is if our enrolment assumptions are incorrect. We believe this to be unlikely; an online MSc in Data Science at the University of Texas accepted 350 students of the over 2000 who applied in its first year of operation. Nonetheless, we have allocated a portion of the proposed budget to advertising and will also screen applicants to our regular Biostatistics MSc and PhD who are unsuitable for those programs (as we are targeting a different audience with a potentially less mathematical background) and offer admission to those students who would be a good fit for the certificates.

We also face risk by not proceeding with this Graduate Certificate. The importance of online teaching – as opposed to the temporary remote learning being undertaken due to the current pandemic – cannot be ignored. By staking a claim to this form of teaching, McGill will reinforce its status as a leader in innovative teaching and will expand its already considerable international reach to student populations who could not otherwise attend our institution. By training the future generations of data scientists ready to deal with important public health challenges, McGill will remain at the forefront of scholarly innovations for the greater good of society.

Appendix: Response to Queries

Market scan

Several programs focusing on data science exist in Canada (Figure 1), most in the form of in-person MSc programs, offering little flexibility to students wishing to remain outside of Canada or remain in full-time employment.

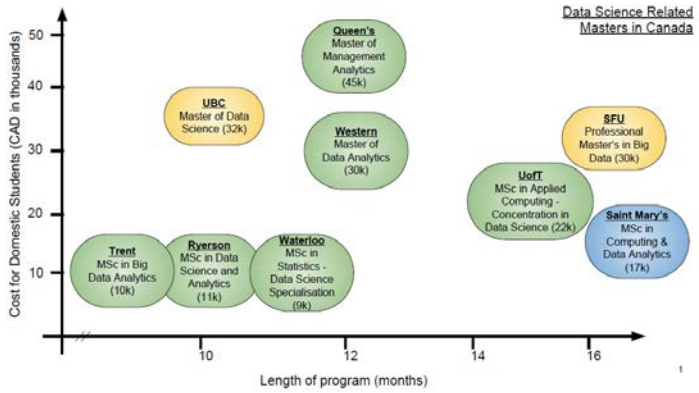

Figure 1. Existing MSc programs in Data Science in Canada (may not be exhaustive).

Interest in our Biostatistics program has increased steadily since its creation as a separate program (formerly, it existed as a "stream" in the Epidemiology program) as shown in Figure 2 (left panel), The increase in applications reflects the increasing demand of (bio)statistics trained graduates in the job market, both in academic and in industry.

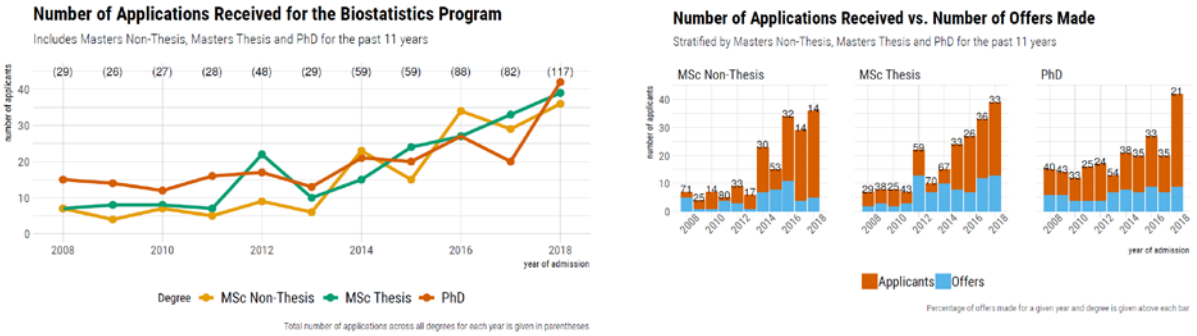

Figure 2. Request for admissions in the Biostatistics program (left) and applications vs. offers made in the Biostatistics program (right).

However, the number of offers made has remained stable (Figure 2, right panel): we have been unable to increase enrolment to meet that need with our current programs for two reasons: (1) many applicants' background is not sufficiently statistical to succeed in our academically-focused programs, and (2) insufficient supervisory capacity in terms of both funding and time – both issues that would not arise with the certificate provided we secure the positions we are requesting.

Through informal discussions with international colleagues as well as those in Canada, we have learned that the University of Glasgow's fully online MSc in Data Science enrols 60-80 students per year. A fully

online MSc in Data Science at the University of Texas – Austin that just opened this year saw over 1000 applicants, and accepted 350 of these into its inaugural cohort. Within Canada, the University of Waterloo's specialization in Data Science within the regular (in-person) Statistics MSc enrolled 60 students last year, while the Data Science program at SFU sees approximately 1000 applicants per year, of whom 10% are accepted (and of those, about 80 students per year take up the offer).

We are therefore extremely confident that we can reach our enrolment targets.

BIOS 624 vs BIOS 643

Concerning the question of whether BIOS 624 could be used in the Graduate Certificate in Health Data Analytics place of the newly proposed BIOS 643: we believe that, in fact, the two are not interchangeable. BIOS 624 is a four-credit course serving students in their second year of the PhD program in Biostatistics as a required course (though also open to other graduate students as an elective), whereas BIOS 643 is to be a three-credit course aimed at students earlier in their training with less exposure to more advanced modelling techniques (e.g. correlated data, missing data methods, and so on). Thus, while some content will be overlapping, the depth and specific topics will differ.