

Programme Information		
Programme Title	Programme Code	HECoS Code
MSc Health Data Analytics and Machine Learning	A3BJ	For Registry Use Only

Award	Length of Study	Mode of Study	Entry Point(s)	Total Credits	
				ECTS	CATS
MSc	1 Calendar Year (12 months)	Full time	Annually in October	90	180
Postgraduate Diploma	N/A	Full time	N/A	60	120
Postgraduate Certificate	N/A	Full time	N/A	30	60

The PG Certificate/PG Diploma are exit awards and are not available for entry. All students must apply to and join the MSc.

Ownership			
Awarding Institution	Imperial College London	Faculty	Faculty of Medicine
Teaching Institution	Imperial College London	Department	School of Public Health
Associateship	N/A	Main Location(s) of Study	St Mary's Campus
External Reference			
Relevant QAA Benchmark Statement(s) and/or other external reference points		N/A	
FHEQ Level		Level 7 – Master's	
EHEA Level		2nd Cycle	
External Accrator(s) (if applicable)			
External Accrator 1:	N/A		
Accreditation received:	N/A	Accreditation renewal:	N/A
Collaborative Provision			
Collaborative partner	Collaboration type	Agreement effective date	Agreement expiry date
N/A	N/A	N/A	N/A
Specification Details			

Programme Lead	Marc Chadeau-Hyam and Paul Elliott
Student cohorts covered by specification	2022-23 entry
Date of introduction of programme	October 18
Date of programme specification/revision	August 22

Programme Overview

The MSc in Health Data Analytics and Machine Learning is a one-year full-time course aimed at building a solid and common scientific expertise in analysing health data. We aim to develop skills in using appropriate cutting-edge quantitative methods to fully exploit complex and high dimensional data that are now emerging in the fields of health sciences. The course will cover mains aspects of

- data handling, storing, and manipulating under strict and rigorous security constraints,
- the in-depth exploration of 'big Medical/Clinical data' using Machine learning approaches
- the analysis and integration of high resolution data including but not restricted to molecular data.

As an emerging scientific field, the course will build upon the analysis of real data (individually or in group) and will ensure that findings from your work first adheres to publication quality standards, and are presented in an interpretable way. To ensure this, you will be taught by experts in the field including international researchers with strong methodological background and expertise in the application of these methods to large-scale medical and clinical data.

In order to establish a common background across multidisciplinary profiles of students, foundational modules will take place in Term 1. These include *Statistical Thinking and Data Analysis* and *Principles and Methods for Epidemiology* which are jointly taught to MSc Epidemiology and MPH students. You will also be taught the HDA-specific modules on (i) Clinical Data Management, (ii) *Molecular Epidemiology* and (iii) *Translational Data Science (Part I)*. In Term 2, the core modules *Translational Data Science Part II*, *Computational Epidemiology and Machine Learning* are taught along with two other shorter (5-weeks) modules, *Advanced Analytics* and *Population Health Analytics*. As students working at a master's level are expected to develop and demonstrate their intellectual independence and are also expected to develop skills for the workplace, Wednesday mornings and afternoons, in both terms, are reserved for optional workshops, seminars and independent study. These include seminar series, student-led journal clubs, additional practicals/tutorial sessions, and student's group presentations in relation to the modules taught in these terms. Term 3 is entirely dedicated to the research project with some additional seminars occurring throughout the term with a focus on developing presentation and communication skills.

The breadth of the content of the course combined with clear orientation towards real-life applications, will offer you a unique experience of the use of these techniques in a real-life context. Building upon the diversity of students' profiles and facilitating cross-fertilisation (e.g. through the constitution of multi-disciplinary groups), we will offer unique teamwork experience. Overall, the strong technical component of the course, combined with its applied orientation, will offer you complementary training that will be essential for the continuation of your career in health data sciences where these skills are in high demand. For instance, after completing this course you will be in an excellent position to:

- 1- pursue an academic career (through a PhD for instance) in many fields: molecular, environmental, chronic disease and computational epidemiology, public health, population sciences;
- 2- work as an expert analyst in industries (pharma, food, cosmetics...) where these big data are also accumulating;
- 3- engage in large data companies or smaller start-up companies which are developing in the field of health and need trained data scientists with experience in exploring and mining medical and epidemiological data experience.

Learning Outcomes

Upon successful completion of the programme, you should be able to:

1. Consider, appraise and describe the in-depth nature, the complexity and the richness of Medical/Clinical/Epidemiological/population-based 'big data';
2. Independently and originally design, maintain, and query novel Medical/Clinical databases, to link and harmonise developed and existing databases and ensure that they comply with legal, ethical, security and privacy standards to ensure a secure use of medical and clinical data;
3. Independently design, implement and interpret results from established statistical models used to perform high-resolution data profiling (univariate approaches, dimensionality reduction techniques, and variable selection);
4. Design, implement, interpret and contextualise results from established machine learning approaches (including decision trees, SVM, random forests, and Bayesian/neural networks);
5. Identify, implement and evaluate the best (set of) statistical and/machine learning method(s) to address a specific research question considering the nature of the data, their specifics and be able to quantitatively compare and differentiate the performances of different possible techniques;
6. Generate original thinking on how to use and combine existing techniques to address in the simplest and most interpretable way pressing questions in the field which involve data integration, results visualisation and interpretation;
7. Formulate original and relevant research questions from existing data sets, construct an analytical plan to address the research questions and appropriately implement the analysis plan leading to an original solution of the research question;
8. Produce a sound piece of scientific work from real data addressing an original research question using appropriate statistical approaches, and developing visualisation resources that ensure a clear and convincing presentation of the work;
9. Debate, discuss and produce collaborative research in a multi-cultural and cross-disciplinary group setting.

The Imperial Graduate Attributes are a set of core competencies which we expect students to achieve through completion of any Imperial College degree programme. The Graduate Attributes are available at: www.imperial.ac.uk/students/academic-support/graduate-attributes

Entry Requirements

Academic Requirement	Normally at least a 2:1 in a science based UK bachelors honours degree or medical degree or equivalent qualification in (maths, statistics, epidemiology biology). For further information on entry requirements, please go to: www.imperial.ac.uk/study/pg/apply/requirements/pgacademic
Non-academic Requirements	N/A
English Language Requirement	IELTS 7 with a minimum of 6.5 in each element or equivalent.
Admissions Test/Interview	Candidates may be invited to participate in an informal online interview in order to assess if the skills and motivations of the candidate are in line with the programme expectations and offering.

The programme's competency standards documents can be found at: <https://www.imperial.ac.uk/study/pg/medicine/health-data-analytics/>

Learning & Teaching Approach

Learning and Teaching Delivery Methods

You will learn through a combination of learning activities that are proven to enable deeper learning and authentic explorations of key skills and experiences. To this end, the learning and teaching activities across modules are organised via a blended approach using asynchronous and synchronous learning methods. These include:

- Pre-session reading
- In class activities such as small group discussions and reflections
- Lectures and associated practicals (synchronous and asynchronous)
- Seminars and technical tutorials
- Small group tutorials
- Group work sessions
- Computer based practical workshops
- Group revision sessions
- Independent learning (pre-reading, homework, consolidation)
- Individual and group research projects and presentations.

Overall Workload

Your overall workload consists of face-to-face and online synchronous sessions, asynchronous sessions and independent learning. While your actual contact hours may vary according to each module, the following gives an indication of how much time you will need to allocate to different activities at each level of the programme. At Imperial, each [ECTS credit](#) taken equates to an expected total study time of 25 hours. Therefore, the expected total study time is 2250 hours per year.

Typically, but this might vary between terms 1 and 2, this would equate to spending around 20% of your time in lectures, practical and seminars (around 400-450 hours) and around 80% of your time (around 1800 – 1850 hours) on independent study, including the research project.

Assessment Strategy

Assessment Methods

The course includes a wide range of summative assessments that include coursework (~20%), practical (~35%) and written examinations (~45%), such as:

- Mid-term and final examinations
- Oral individual and group presentations
- Online quizzes
- Short answer essay questions
- Individual and group real case analyses reports
- Final research paper/dissertation.

These are split across the terms as follows:

Written Examinations (45%)	End of Term 1 modules (January) End of Term 2 modules (April/May)
Coursework Assessments (55%)	Continuous
Project Deadlines	Continuous (mini group projects) End of Term 1 modules (January) End of Term 2 modules (April/May) Term 3 research project (First week of September)
Practical Assessments	Continuous Term3 Research project viva (Mid/late September)

The wide range of summative assessments aims to test if you have gained the relevant technical, practical, critical thinking and presentation skills that are required to successfully complete the course and start your career in the area of your choice. This includes, for example, assessing if you have: acquired the terminology used in the area; the ability to critically think about a problem and solve it by considering original and relevant research questions from existing data sets and using appropriate techniques to address them; the capacity to independently design, implement and interpret results from established and cutting-edge statistical models used to perform high throughput data profiling (univariate approaches, dimensionality reduction techniques, and variable selection) and integration; demonstrated a systematic understanding, knowledge and expertise in statistical and machine learning approaches and their utility in medical data analytics and to be able to design,

implement and interpret results from such approaches; and, being able to work in a team with other students who have different backgrounds and viewpoints.

Academic Feedback Policy

The programme complies with the College policy and will provide students with feedback within a timely and appropriate turnaround time. Students will be provided with a bespoke assessment schedule confirming all submission deadlines, marking periods and feedback points to manage their expectations and ensure feedback can inform the next assessment.

Feedback will take different forms according to the type of assessment:

- Exams and written reports: students could be given written feedback, and given the opportunity to discuss their provisional marks and possible improvements with the module leaders;
- Oral presentations: comments and immediate feedback will be given during the presentation and on the evaluation sheet; further discussion could be organised once provisional marks are communicated to students;
- Final research project: students will be given written feedback for their final project/presentation and also receive oral feedback following the oral presentation.

Further exchanges will also take place throughout the programme through:

- immediate feedback through online mini-quizzes;
- questions during or after the lectures;
- interaction with tutors during practicals, group work supervisors and 1-2-1s.

Re-sit Policy

The College's Policy on Re-sits is available at: www.imperial.ac.uk/student-records-and-data/for-current-students/undergraduate-and-taught-postgraduate/exams-assessments-and-regulations/

Mitigating Circumstances Policy

The College's Policy on Mitigating Circumstances is available at: www.imperial.ac.uk/student-records-and-data/for-current-students/undergraduate-and-taught-postgraduate/exams-assessments-and-regulations/

Additional Programme Costs Not applicable. Can we just remove it?

This section should outline any additional costs relevant to this programme which are not included in students' tuition fees.

Description	Mandatory/Optional	Approximate cost
A computing machine is required for this course.	Mandatory	Not Provided

Important notice: The Programme Specifications are the result of a large curriculum and pedagogy reform implemented by the Department and supported by the Learning and Teaching Strategy of Imperial College London. The modules, structure and assessments presented in this Programme Specification are correct at time of publication but might change as a result of student and staff feedback and the introduction of new or innovative approaches to teaching and learning. You will be consulted and notified in a timely manner of any changes to this document.

Programme Structure ¹					
FHEQ Level 7 Students study all compulsory and core modules.					
Code	Module Title	Core/ Elective/ Compulsory	Group	Term	Credits
TBC	Introduction to Statistical Thinking and Data Analysis	Compulsory	7	1	7.5
PUBH70026	Principles and Methods of Molecular Epidemiology	Compulsory	7	1	7.5
TBC	Molecular Epidemiology	Compulsory	7	1	7.5
PUBH70028	Translational Data Science Part I and Part II	Core	7	1,2	5
PUBH70029	Clinical Data Management	Compulsory	7	1	7.5
TBC	Machine Learning	Core	7	2	7.5
PUBH70031	Computational Epidemiology	Core	7	2	7.5
PUBH70032	Advanced Analytics	Compulsory	7	2	5
PUBH70033	Population Health Analytics	Compulsory	7	2	5
PUBH70034	Research Project	Core	7	3	30
Credit Total					90

¹ **Core** modules are those which serve a fundamental role within the curriculum, and for which achievement of the credits for that module is essential for the achievement of the target award. Core modules must therefore be taken and passed in order to achieve that named award. **Compulsory** modules are those which are designated as necessary to be taken as part of the programme syllabus. Compulsory modules can be compensated. **Elective** modules are those which are in the same subject area as the field of study and are offered to students in order to offer an element of choice in the curriculum and from which students are able to select. Elective modules can be compensated.

Progression and Classification

Upon completion of the course, the MSc can be awarded as a Pass, Merit or Distinction.

Award and Classification for Postgraduate Students

Award of a Postgraduate Certificate (PG Cert)

To qualify for the award of a postgraduate certificate a student must have a minimum of 30 credits at Level 7 (this may include a maximum of 10 credits from Level 6 where this is approved as part of the award).

Award of a Postgraduate Diploma (PG Dip)

To qualify for the award of a postgraduate diploma a student must have passed modules to the value of no fewer than 60 credits at Level 7 (this may include a maximum of 15 credits from Level 6 where this is approved as part of the award).

1. and no more than 10 credits as a Compensated Pass;

Classification of Postgraduate Taught Awards

The College sets the class of Degree that may be awarded as follows:

1. Distinction: The student has achieved an overall weighted average of 70.00% or above across the programme.
2. Merit: The student has achieved an overall weighted average of above 60.00% but less than 70.00%.
3. Pass: The student has achieved an overall weighted average of 50.00% but less than 60.00%.
 - a. For a Masters, students must normally achieve a distinction (70.00%) mark in the dissertation or designated final major project (as designated in the programme specification) in order to be awarded a distinction.
 - b. For a Masters, students must normally achieve a minimum of a merit (60.00%) mark in the dissertation or designated final major project (as designated in the programme specification) in order to be awarded a merit
 - c. Modules taken at level 6 as part of the programme specification for a named postgraduate award will contribute to the determination of pass, merit or distinction for any taught postgraduate award and are included in the calculation of the overall weighted average.

Programme Specific Regulations

N/A

Supporting Information

The Programme Handbook is available at: <https://www.imperial.ac.uk/study/pg/medicine/health-data-analytics/>

The Module Handbook is available at: **TBA**

The College's entry requirements for postgraduate programmes can be found at:
www.imperial.ac.uk/study/pg/apply/requirements

The College's Quality & Enhancement Framework is available at:
www.imperial.ac.uk/registry/proceduresandregulations/qualityassurance

The College's Academic and Examination Regulations can be found at:
www.imperial.ac.uk/about/governance/academic-governance/regulations

Imperial College is an independent corporation whose legal status derives from a Royal Charter granted under Letters Patent in 1907. In 2007 a Supplemental Charter and Statutes was granted by HM Queen Elizabeth II. This Supplemental Charter, which came into force on the date of the College's Centenary, 8th July 2007, established the College as a University with the name and style of "The Imperial College of Science, Technology and Medicine".
www.imperial.ac.uk/admin-services/secretariat/college-governance/charters/

Imperial College London is regulated by the Office for Students (OfS)
www.officeforstudents.org.uk/advice-and-guidance/the-register/

This document provides a definitive record of the main features of the programme and the learning outcomes that a typical student may reasonably be expected to achieve and demonstrate if s/he takes full advantage of the learning opportunities provided. This programme specification is primarily intended as a reference point for prospective and current students, academic and support staff involved in delivering the programme and enabling student development and achievement, for its assessment by internal and external examiners, and in subsequent monitoring and review.

Modifications

Description	Approved	Date	Paper Reference
Curriculum Review	Programmes Committee	29/03/22	PC.2021.86