PROGRAMME SPECIFICATION

1. Overview

Programme Title	MSc Health Data Science
Academic Year	2027-28
Cohort Entry Points	Annually in September
Programme Director	Stéphane Hué
Awarding Body	London School of Hygiene & Tropical Medicine
Teaching Institution	London School of Hygiene & Tropical Medicine
Faculty	Epidemiology and Population Health
Length of Programme	Full-time Taught Masters programmes: 12 months Part time/split study Taught Masters programmes: 24 months
Mode of Delivery	Intensive
Maximum Period of Registration	Full-time 3 years/Part-time 5 years
Language of Study	English
Entry Requirements	Please refer to the specific programme entry requirements here.
Award Titles	Master of Health Data Science Exit Awards: Postgraduate Diploma in Health Data Science (120 credits) Postgraduate Certificate in Health Data Science (60 credits)
Total Credits	**CATS:** 180 **ECTS:** 90
Accreditation by Professional Statutory and Regulatory Body	None

Relevant PGT QAA Benchmark Statement and/or other External/Internal Reference Points	Consistent with the Framework for Higher Education Qualifications at Masters level (Level 7), this programme will provide students with an understanding of key concepts relevant to health data science and training in essential tools and skills to manage and analyse very large diverse datasets across healthcare systems and to communicate results appropriately to diverse audiences.
Framework for Higher Education Qualifications Level (FHEQ)	Level 7
Re-sit Policy	https://www.lshtm.ac.uk/sites/default/files/academic-manual-chapter-08a.pdf
Extenuating Circumstances Policy	https://www.lshtm.ac.uk/sites/default/files/academic-manual-chapter-07.pdf
Date of Introduction of Programme	October 2020

2. Programme Description, Aims & Learning Outcomes

Programme Description
Health Data Science is an emerging discipline, combining mathematics, statistics, epidemiology and informatics. This programme will equip graduates with the tools and skills to manage and analyse very large diverse datasets across healthcare systems. This programme aims to train a new generation of world-leading health data scientists, to work in both the public and private sector. The overall strategy towards teaching and assessment focuses on building strong quantitative, computational, and practical data management skills, while providing opportunities to develop key professional skills required to be a successful health data scientist.
Educational Aims of the Programme
The programme aims are to: • equip graduates with the tools and skills to manage and analyse very large diverse datasets across healthcare systems;

- provide opportunities to allow students to develop the professional skills – including teamwork, project management, and presentation skills – to work as a successful data scientist in the public or private sector.

Programme Learning Outcomes

1. Knowledge and Understanding of:

- the varied roles of a health data scientist within the wider health and clinical research environment;
- key sources of health data (e.g. clinical, social, genomic, imaging, digital, environmental) and their collection);
- health data quality, governance and accessibility, including how, including how the context of data collection influences bias and appropriate use;
- commonly used statistical, epidemiological, machine learning and artificial intelligence techniques, and their appropriate application to health problems;
- key issues related to ethics, security and information governance in the current context of health data science and responsible AI.

2. Skills and Other Attributes

Intellectual Skills – able to:

- design and implement a rigorous analysis plan using appropriate statistical, machine learning, and AI techniques to address a defined health research question;
- critically appraise ethical, security and information governance implications of a proposed study design in the context of a data science project;
- identify and mitigate potential sources of bias, assessing their likely impact on inference and decision-making;
- analyse and interpret results responsibly and effectively, considering implications for practice, policy, or patient outcomes.

Practical Skills – able to:

- extract, assemble, clean, and manage diverse health datasets (e.g. electronic health records, registries, genomic, or digital data) within a reproducible workflow;
- implement data pipelines and analytical solutions using appropriate tools (e.g., R, Python, SQL, and APIs) in line with open-science and FAIR data principles.

Transferable Skills – able to:

- Engage with non-technical stakeholders or clients to understand and agree project scope, effectively manage a data science project, and deliver key objectives within a set timescale;
- work effectively within a multi-disciplinary environment, both independently and as an effective team member;
- communicate complex technical methods and results to a mixed audience through written reports and oral presentations;
- reflect on the broader societal and environmental impact of health data science, demonstrating integrity, inclusivity, and commitment to public benefit.

Teaching and Learning Strategy

The role of the health data scientist requires a wide range of technical, practical, and professional skills, many of which are best developed by hands-on experience grappling with real-world problems. Professional skills, including project management, communicating to diverse audiences, and effective teamwork, are also crucial. The teaching approach for this MSc is designed to maximise students' time working on practical problems, in individual and group settings, and will require students to interact with a range of collaborators/clients.

As well as traditional lectures followed by problem-based practical sessions, with or without computers, teaching strategies in the programme will include:

- **Flipped classroom** approaches where students are provided with materials to read/watch independently, followed by formative assessment in class to assess understanding (e.g. via Moodle-based multiple choice questions), allowing contact time to focus on practical problem-based learning.
- **Interactive lectorials (learning and teaching sessions that combine elements of both a formal lecture and an interactive seminar)** alternating lecture-based and hands-on practical sessions.
- **Panel discussions and workshops**, to stimulate debate particularly for current live controversies such as the ethics of algorithms.
- **Teamwork**, particularly in the team-based module and the hackathon.
- **Opportunities to develop and practice professional skills**, including a range of student-led presentations, modules which require student teams to interact with a client (someone who is not a data scientist working outside of the LSHTM who wishes to "employ" our students to address a particular research question).

Assessment Strategy
Assessments have been designed to reflect the reality of life as a health data scientist. The programme will include a mix of formative and summative assessment. A range of assessment techniques will be deployed, including: • In-module examinations • Submission of code to perform a given task • Oral presentations of technical material • Coursework, with structured and unstructured questions • Comprehensive written report of an in-depth exploration into an area chosen by the student (research project)

3. Programme Structure and Features, Modules, Credit Assignment and Award Requirements:

Full-time Masters	**Term 1**	**Term 2/3**	**Term 3**	**Total Credits**
Compulsory Modules	5	2	0	90
Recommended Modules	0	2	0	30
Project	0	0	1	60

Module information is correct at the time of publication, but minor amendments may be made subject to approval as detailed in Chapter 3 of the LSHTM Academic Manual. Elective modules listed are indicative and may change from year to year. https://www.lshtm.ac.uk/study/courses/changes-courses

Term	**Slot**	**Module Code**	**Module Title**	**Module Type**	**Credits (CATS)**
1	AB1	2485	Thinking Like a Health Data Scientist	Compulsory	10
1	AB1	2486	Programming	Compulsory	10
1	AB1	2487	Health Data Management	Compulsory	15
1	AB1	2488	Concepts and Methods in Epidemiology	Compulsory	10
1	AB1	2489	Statistics for Health Data Science	Compulsory	15
2	C1	2490	Machine Learning	Compulsory	15
2	C2	2491	Data Challenge	Compulsory	15
2	D1	2492	Genomics Health Data	Recommended	15

2	D1	2465	Analysis of Hierarchical and Other Dependent Data	Recommended (Need to request permission from module organisers)	15
2	D1	3135	Spatial epidemiology in Public Health	Recommended	15
2	D1	2611	Ai in Health: Principles and Applications	Recommended	15
2	D2	2496	Bayesian Analysis	Recommended (Need to request permission from module organisers)	15
2	D2	2493	Analysis of Electronic Health Record Data	Recommended	15
2	D2	1301	Environmental Epidemiology	Recommended	15
3	E1	2464	Modelling & the Dynamics of Infectious Diseases	Recommended	15

MSc Health Data Science students wishing to take the module Modelling & the Dynamics of Infectious Diseases (2464) must take three modules in the four teaching blocks in Term 2, leaving one block from the four (C1, 2, D1, D2) free for Project preparation. During all these four blocks some (non-assessed) support will be available to help with the planning of student Projects.

Contact Time

Student contact time refers to the tutor-mediated time allocated to teaching, provision of guidance and feedback to students. This time includes activities that take place in face-to-face contexts such as on-campus lectures, seminars, demonstrations, tutorials, supervised laboratory workshops, practical classes, project supervision and external fieldwork or visits, as well as where tutors are available for one-to-one discussions and interaction by email. Module contact time will be defined in the individual module specifications and provided to students at the start of their programme.

This definition is based on the one provided by the Quality Assurance Agency for Higher Education (QAA) Explaining contact hours (2011). Student contact time, together with time allocated for independent study and assessment, determines the total student study hours for a module or programme. Although there are separate hours allocated for each of these activities, they should always be clearly linked together to support effective learning.

The London School of Hygiene and Tropical Medicine (LSHTM) defines high quality contact time as structured, focused, purposeful and interactive.