Tutor notes for Clinical Bioinformatics I

A part of the MSc in Clinical Science (Clinical Bioinformatics)

Covering:

- course content,
- intended learning outcomes,
- assessment processes
- example scenarios

University of Manchester

Version history:

v1.0 created 27th September 2013 by Andy Brass
Section 1: Course content and syllabus

Background

These notes relate to the 10 credit section of this module relating to the introduction to clinical bioinformatics.

Students will be working in small groups (4-5 per group) that have been allocated during the induction week.

Each group has been allocated a scenario (see section 3 for examples) that they will be working on during the week.

On each of the 5 days of the course they will be receiving lectures in the morning and then will work on a specific part of the scenario in the afternoon.

The afternoon sessions will be divided into two tasks. These short tasks are all related and will combine together to provide a framework to allow them to successfully complete the group exercise (scenario analysis).

Each subtask will have specific learning objectives associated with it as listed in this document. It is your job as a tutor to check that the students are working towards these objectives. If all is going well, then you can let the students get on with things. However, if you feel that they are exploring blind alleys, getting confused, misinterpreting the task, then please nudge them on track with a few well-chosen question (“do you feel this approach is appropriate..”, “have you perhaps thought of considering ...?”). At the end of each section you need to be comfortable that your group has covered the learning objectives of each session.

At the end of each task every group will be expected to provide a short report on what they have achieved. As tutors we will provide formative feedback on these reports after all the groups have contributed. If any groups have missed some of their learning objective despite a tutors best efforts, this can be highlighted at this point in the afternoon.

As a tutor you are also expected observe your group carefully so as to be able to complete a final summative assessment of the group performance. The marking scheme used for this assessment is described in section 2.
Day 1: The genetics and biology of disease

Material covered in the morning lectures:

**Genetics/Genomics**

- Introduction to the history and scope of genomics
- The Genome Landscape
- Nucleic Acid structure and function, including the structure and function of coding and non-coding DNA
- The central dogma
- From DNA, to RNA and proteins
- Noncoding regulatory sequence: promoters, transcription factor binding sites, splice site dinucleotides, enhancers, insulators
- Genetic variation and its role in health and disease
- Genomic technology and role of the genome in the development and treatment of disease

**Sequencing**

- Types of sequencing, applications and limitations; Sanger versus short read
- Analysis, annotation and interpretation
- Panel versus exome versus whole genome resequencing

**Statistics**

- Basic statistics applied to clinical genetics/genomics
- Hardy-Weinberg, Bayes theorem, risks in pedigrees

The afternoon session: exploring the scientific literature

Each group has been given a scenario. On day 1 of the scenario a patient has been referred for genetic testing. At this stage there is a suspicion as to why there might be a concern, but no results have been obtained. The aim of this session is for the trainees to develop good skill sets for retrieving information around the genetics of the disease of interest, to develop standard operating protocols (SOPS) that capture best practice in the area, and finally to retrieve key facts from the retrieved literature discussing genotype/phenotype relationships for the disease of interest.

**Part 1: Locating the information**

Key learning objectives: strategies for effective and clinically relevant information retrieval from scientific and other resources.

- To explore what a scientific paper is as an artefact, how is it constructed (useful information could be in methods or results sections, some data could be in supplementary information, the refereeing process, the importance of citations).
- To understand where scientific literature is stored (journals/online).
• To gain an understanding of tools for retrieving papers – in particular systems such as scholar google for tracking citations.
• To develop effective online-search strategies and the importance of capturing the process in an SOP.
• To explore non-refereed sources – patient blogs, generic internet, curated internet resources (Omim).
• To gain an understanding of tools for managing references.

Short report: Each group to give a short informal presentation on the strategy (SOP) they are using to find the relevant papers and information.

Formative feedback from tutors

Part 2: Extracting clinically relevant meaning:

Key learning objectives: The production of concise, informative and authoritative statements based on a review of literature

• To be able to extract key facts from the literature – what is it that the paper is claiming, is it believable, how would you know?
• To provide evidence of authority (use of citation counts, journal impact factors.
• To be able to illustrate good referencing to provide audit.

Short report: Each group to put together a short document highlighting the key facts (bullet points are acceptable, but must have good citations and evidence for source authority)

Formative feedback from tutors
Day 2: Basic bioinformatics

Material covered in the morning lectures:

**Bioinformatic Fundamentals**

- Introduction to the history and scope of bioinformatics
- Primary biological sequence resources, including INDSC (GenBank, EMBL, DDBJ) and UniProt (SwissProt and TrEMBL)
- Genome browsers and interfaces; including Ensembl, UCSC Genome Browser, Entrez, Similarity/homology, theory of sequence analysis, scoring matrices, dynamic programming methods including BLAST, pairwise alignments (e.g., Smith Waterman, Needleman Wunsch), multiple sequence alignments (e.g., ClustalW, T-Coffee, Muscle), BLAT
- Feature identification including SNP analysis and transcription factor binding sites and their associated TF binding sequence motifs
- Ontologies – in particular GO, Human Phenotype Ontology (HPO)

The afternoon session: Basic bioinformatics strategies

Each group has been given a scenario to explore. On day 2 the results of the genetics test have been obtained and a variant has been identified. The aim of this session is for the trainees to develop good skill sets for exploring what the consequences of this variation might be, to develop standard operating protocols (SOPS) that capture best practice in the area, and finally to run a number of bioinformatics analyses from which a report discussing the consequences will be constructed.

**Part 1: A bioinformatics strategy**

Key learning objectives: strategies for effective analysis of single gene variants using bioinformatics tools.

- To explore what bioinformatics resources it might be appropriate to use for the analysis.
- To explore what bioinformatics tools might be appropriate to use for the analysis.
- To critically evaluate resources for their applicability and authority.
- To understand the importance of combining multiple analyses to look for concordance/discordance.
- The development of a bioinformatics analysis pipeline to determine the consequences of the variation.

**Short report:** Each group to give a short informal presentation on the strategy (SOP) they are using to analyse the variant.

**Formative feedback from tutors**

**Part 2: A bioinformatics interpretation:**

Key learning objectives: The production of concise, informative and authoritative statements based on an analysis of the bioinformatics results

- To performing a number bioinformatics analyses based on the variations they have been given
To interpret the data returned effectively.
To critically evaluate the data.
To integrate of information from different sources.
To understand the importance of capturing an analysis pipeline in an SOP or report.

**Short report:** Each group to put together a short document highlighting the key facts obtained from the analyses (bullet points are acceptable, but must have evidential support).

**Formative feedback from tutors**
Day 2: Clinical bioinformatics

Material covered in the morning lectures:

**Clinical application of bioinformatics**

Introduction to the clinical application of bioinformatic resources, including its role and use in a medical context in molecular genetics, cytogenetics and next generation sequencing for data manipulation and analysis, and genotyping microarrays (also used to predict CNVs).

- Background and application of specialist databases and browsers
- dbSNP
- DECIPHER
- Orphanet
- DMuDB / NGRL Universal Browser
- OMIM
- ECARUCA
- DGV
- LOVD/UMD database software and scientific literature
- HGMD
- Specific clinical analysis software
- CNV analysis
- Gene Prioritisation (e.g. ToppGene, Endeavour, GeCCO)
- Missense analysis (e.g. Align GVGD, SIFT, PolyPhen, Panther, PhDSNP, MAPP)
- Splicing analysis applications (e.g. GeneSplicer, MAxEntScan, NNSplice, SSFL, HSF, NetGene2)
- Commercially available software (e.g. NextGENe, Alamut, Cartegenia
- Capture and representation of phenotype data
- Development of a simple application for clinical bioinformatic use

The afternoon session: Clinical bioinformatics strategies

Each group has been given a scenario to explore. On day 3 the trainees will place what they developed and learnt on day 2 into a specifically clinical context. The aim of this session is for the trainees to develop good skill sets for exploring what the consequences of this variation might be in the context of existing knowledge around role of variation in health disease. Not all the variations seen will have a clinical consequence, not all the variations seen will have a known or predictable phenotype.

**Part 1: A clinical bioinformatics strategy**

Key learning objectives: strategies for effective analysis of single gene variants in the context of clinical bioinformatics data.

- To explore what clinical bioinformatic resources might be appropriate to use for the analysis.
- To explore what clinical bioinformatic tools might be appropriate to use for the analysis.
• To develop an analysis pipeline to determine the consequences of the variation supported by existing clinical bioinformatics knowledge.
• To understand the importance of data standards in working with genetics data in the context of the NHS.
• To explore some of the possible governance and privacy issues that such analyses might trigger.

**Short report:** Each group to give a short presentation on the strategy (SOP) they are using to analyse the variant – and any potential and privacy issues that such an analysis might cause.

**Formative feedback from tutors**

**Part 2: Clinical interpretation:**

Key learning objectives: The production of concise, informative and authoritative statements based on an analysis of the clinical bioinformatics results.

• To perform a number clinical bioinformatics analyses based on the variations they have been given.
• To interpret the data returned effectively.
• To critically evaluate the data.
• To Integrate information from different sources.
• To understand the importance of capturing an analysis pipeline in an SOP or report.

**Short report:** Each group to put together a short document highlighting the key facts obtained from the analyses (bullet points are acceptable, but must have evidential support)

**Formative feedback from tutors**
Day 4: Case studies

**Material covered in the morning session:**

Case studies and tutorials from clinical scientists from the NHS working in clinical genetics and bioinformatics. These will cover the practical aspects of the role of a clinical bioinformatician in the NHS, particularly focusing on areas such as governance and standards. There will also be talks that put the role of the clinical in in the wider clinical context of the patient pathway and the part they play in the clinical decision making process.

**The afternoon session: Clinical bioinformatics strategies**

In this session the students will continue the analyses based on their scenarios. This will provide students and tutors a chance to develop a deeper analysis of each of their scenarios – particularly in the light of the information received regarding the clinical implications of the clinical bioinformatician’s work.

Key learning objectives: The production of concise, informative and authoritative statements based on an analysis of the clinical bioinformatics results that are of a sufficient standard to provide input into a multidisciplinary team meeting.

**Short report:** Each group to put together a short report summarising the key findings of the week in a form suitable for use in a multidisciplinary team meeting.

**Formative feedback from tutors** – and in particular areas that the groups need to focus on to complete a successful scenario.
Day 5: Completion of the scenario

Morning session

Groups to complete their analyses and prepare a report suitable for submission to a clinical care team.

The afternoon session:

Each group will provide an oral report on their scenario and how they believe their findings could be used to support future treatment decisions for the patient.

**Formative feedback from a clinical scientist:** A practicing clinical scientist will provide feedback on the presentations from the groups – focussing in particular on the analysis strategies followed, and the appropriateness of the report produced given its target clinical audience.

Final activity: Participants to be handed the scenarios they will work on individually for their module assessment.
Tutors notes: Clinical Bioinformatics 1

Section 2: Assessment Procedures

Background

These notes relate to the 10 credit section of this module relating to the introduction to clinical bioinformatics.
Students will be working in small groups (4-5 per group) that have been allocated during the induction week.
Each group has been allocated a scenario (see attached sheet for examples) that they will be working on during the week.
This document lists the assessments that will be applied in this module.

Formative feedback:

"it can be argued that giving learners feedback is just about the most important dimension of the work of teachers in post-compulsory education ... but perhaps all told, formative feedback is the vital dimension, as given at the right time and in the best possible way it can lead learners steadily towards successful achievement in summative assessment contexts." Race P. (2005) “Making Learning Happen”, Chapter 5, Learning through Feedback, London, Sage

On every day of the course students will present on the progress they have made on that day’s specific task. As a tutor you will need to provide feedback on this presentation. It is important that as well as highlighting areas of strength and weakness in the students work, you also provide information on what they could do to help improve any weaknesses.

Summative assessment

This is the assessment provided to measure student performance. There are two aspects of the student work that will be assessed – their performance in the group work, and the report they produce on the individual scenario they are given on day 5. Mark sheets and assessment criteria for these are provided on the following pages. Assessment 1 (marked out of 30) contributes 40% of the mark, Assessment 2 (marked out of 40) contributes 60% of the marks.

The final mark for these 10 credits:

The calculation to return a % mark for these 10 credits is therefore:

Final mark (100) = (Assessment 1 mark/30)*40 + (Assessment 2 mark/40) * 60
## INTENDED LEARNING OUTCOMES (ILOs) – what the assessments are testing

<table>
<thead>
<tr>
<th>Category of outcome</th>
<th>Introduction to Clinical Bioinformatics and Genomics</th>
</tr>
</thead>
</table>
| Knowledge and understanding | 1. Discuss the governance and ethical frameworks in place within the NHS and how they apply to bioinformatics.  
2. Discuss and justify the importance of standards, best practice guidelines and standard operating procedures: how they are developed, improved and applied to clinical bioinformatics.  
3. Describe the structure of DNA and the functions of coding and non-coding DNA.  
4. Discuss the flow of information from DNA to RNA to protein in the cell.  
5. Describe transcription of DNA to mRNA and the protein synthesis process.  
6. Discuss the role of polymorphisms in Mendelian and complex disorders and give examples of polymorphisms involved in genetic disease.  
7. Describe appropriate bioinformatics databases capturing information on DNA, RNA and protein sequences.  
8. Explain the theory of sequence analysis and the use of genome analysis tools.  
10. Explain fundamental bioinformatic principles, including the scope and aims of bioinformatics and its development.  
11. Explain fundamental genomic principles, including the scope and aims of genomics and its development.  
12. Discover resources linking polymorphism to disease processes and discuss and evaluate the resources that are available to the bioinformatician and how these are categorised.  
13. Discuss metadata and how it is captured in bioinformatics resources.  
14. Interpret the metadata provided by the major bioinformatics resources.  
15. Describe the use of ontologies in metadata capture and give examples of the use of ontologies for capturing information on gene function and phenotype.  
16. Identify appropriate references where published data are to be reported.  
17. Describe the biological background to diagnostic genetic testing and clinical genetics, and the role of bioinformatics.  
18. Describe the partnership of Clinical Bioinformatics and Genetics to other clinical specialisms in the investigation and management of genetic disorders and the contribution to safe and effective patient care. |

### Intellectual skills

<table>
<thead>
<tr>
<th>Intellectual skills</th>
<th>These ILOs tested in the report</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Critically analyse scientific and clinical data</td>
<td></td>
</tr>
<tr>
<td>2. Present scientific and clinical data appropriately</td>
<td></td>
</tr>
<tr>
<td>3. Formulate a critical argument</td>
<td></td>
</tr>
<tr>
<td>4. Evaluate scientific and clinical literature</td>
<td></td>
</tr>
<tr>
<td>5. Apply the knowledge of clinical bioinformatics to address specific clinical problems</td>
<td></td>
</tr>
</tbody>
</table>

### Practical skills

<table>
<thead>
<tr>
<th>Practical skills</th>
<th>These ILOs tested in the group work</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Present information clearly in the form of verbal and written reports.</td>
<td></td>
</tr>
<tr>
<td>2. Communicate complex ideas and arguments in a clear and concise and effective manner.</td>
<td></td>
</tr>
<tr>
<td>3. Work effectively as an individual or part of a team.</td>
<td></td>
</tr>
<tr>
<td>4. Use conventional and electronic resources to collect, select and organise complex scientific information</td>
<td></td>
</tr>
<tr>
<td>5. Perform analysis on DNA data and protein sequence data to infer function.</td>
<td></td>
</tr>
<tr>
<td>6. Perform sequence alignment tasks.</td>
<td></td>
</tr>
<tr>
<td>7. Select and apply appropriate bioinformatic tools and resources from a core subset to typical diagnostic laboratory cases, contextualised to the scope and practice of a clinical genetics laboratory.</td>
<td></td>
</tr>
<tr>
<td>8. Compare major bioinformatics resources for clinical diagnostics, and how their results can be summarised and integrated with other lines of evidence to produce</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>9.</td>
<td>Interpret evidence from bioinformatic tools and resources and integrate this into the sum of genetic information for the interpretation and reporting of test results from patients.</td>
</tr>
<tr>
<td>10.</td>
<td>Perform the recording of building or version numbers of resources used on a given date, including those of linked data sources, and understand the clinical relevance of this data.</td>
</tr>
<tr>
<td><strong>Transferable skills and personal qualities</strong></td>
<td><strong>These ILOs tested in the group work and the final report</strong></td>
</tr>
<tr>
<td>1.</td>
<td>Present complex ideas in simple terms in both oral and written formats.</td>
</tr>
<tr>
<td>2.</td>
<td>Consistently operate within sphere of personal competence and level of authority.</td>
</tr>
<tr>
<td>3.</td>
<td>Manage personal workload and objectives to achieve quality of care.</td>
</tr>
<tr>
<td>4.</td>
<td>Actively seek accurate and validated information from all available sources.</td>
</tr>
<tr>
<td>5.</td>
<td>Select and apply appropriate analysis or assessment techniques and tools.</td>
</tr>
<tr>
<td>6.</td>
<td>Evaluate a wide range of data to assist with judgements and decision making.</td>
</tr>
<tr>
<td>7.</td>
<td>Interpret data and convert into knowledge for use in the clinical context of individual and groups of patients.</td>
</tr>
<tr>
<td>8.</td>
<td>Work in partnership with colleagues, other professionals, patients and their carers to maximise patient care.</td>
</tr>
</tbody>
</table>
Mark Sheet for Assessment 1: Group work performance

This assessment is designed to examine a trainee’s contribution to the group work. There are three parts to the mark awarded. The first part of the mark reflects how effectively the group has functioned, (Marked out of 10). The second part reflects the contribution an individual has made to the group (mark out of 10). The third part reflects the quality of the presentations made by the group on the scenario sub-sections. (Mark out of 10). This assessment is testing the trainee’s performance against many of the practical and personal and transferable skill learning outcomes. To score a mark of 5 or greater for a component you need to be clear that there is evidence to support some degree of attainment in all outcomes. This mark contributes 40% of the mark awarded for this 10 credit subsection of Clinical Bioinformatics 1.

<table>
<thead>
<tr>
<th>Student name:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Marker name:</td>
<td></td>
</tr>
<tr>
<td>Date:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group effectiveness (1-10):</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Comments</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Individual contribution (1-10):</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Comments</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quality of presentations (1-10):</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Comments</td>
<td></td>
</tr>
</tbody>
</table>

Final mark (30)
Assessment criteria:

<table>
<thead>
<tr>
<th>Section</th>
<th>0-4</th>
<th>5-6</th>
<th>6-7</th>
<th>7-10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group Effectiveness</strong></td>
<td>Group very disorganised. No responsibility taken for assigning tasks. Unprepared for most of the informal presentations of group work.</td>
<td>Group functions effectively. Roles are sometimes assigned and there is evidence that the some work is distributed effectively amongst team members. Team is prepared for the presentations.</td>
<td>Group runs well with roles assigned and tasks are well distributed amongst team members.</td>
<td>A team working to professional standards. Work distribution and communication is excellent, and builds and supports individual member strengths and development needs.</td>
</tr>
<tr>
<td><strong>Individual contribution</strong></td>
<td>No evidence of communication with other group members, or behaviour towards rest of group disruptive. No input into group objectives.</td>
<td>Evidence for communication with other group members and positive contribution to group discussions. Effective input into group objectives.</td>
<td>Good understanding of group roles and strength and weaknesses of team members. Aware of needs of other group members and helps in their development. Important contributions to group tasks.</td>
<td>Provides constructive guidance and can use knowledge of group strengths and weaknesses to modify their own behaviour to optimise group performance. Good feedback provided to others on the contributions they have made.</td>
</tr>
<tr>
<td><strong>Quality of presentations</strong></td>
<td>Minimal content and mostly missing key targets.</td>
<td>Presentations cover the minimum required learning objectives for the unit.</td>
<td>Presentations show a good understanding of the topics and in a number of cases exceeds the minimum learning objectives.</td>
<td>Excellent presentations showed a good and detailed knowledge, and in most cases exceeded the expected learning objectives.</td>
</tr>
</tbody>
</table>
Mark Sheet for Assessment 2: Individual scenario report

This assessment is designed to examine a trainee's ability to work through the scenario they have been assigned. This will be presented as a 3000 word report. There are four parts to the mark awarded corresponding to students' ability to develop an understanding of the disease area, the appropriateness of the analysis performed, their ability to summarise this information in a way relevant to a clinical scientist, and the general professionalism of the presentation. This assessment is testing knowledge and understanding and intellectual skills learning outcomes. To score a mark of 5 or greater for a component you need to be clear that there is evidence to support some degree of attainment in all outcomes. This mark contributes 60% of the mark awarded for this 10 credit subsection of Clinical Bioinformatics 1.

<table>
<thead>
<tr>
<th>Student name:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Marker name:</td>
<td></td>
</tr>
<tr>
<td>Date:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario background research (10):</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Comments</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bioinformatics strategy (10):</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Comments</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Medical consequences (10):</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Comments</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Report Quality (10)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Comments</td>
<td></td>
</tr>
</tbody>
</table>

Final mark (40)
Assessment criteria:

<table>
<thead>
<tr>
<th>Section</th>
<th>0-4</th>
<th>5-6</th>
<th>6-7</th>
<th>7-10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scenario background research</strong></td>
<td>Includes minimal relevant information – contains many errors or much of what is included is irrelevant to the topic</td>
<td>Covers several issues, but these are listed rather than linked.</td>
<td>Several issues are discussed and related to the problem.</td>
<td>Covers all the major areas, places them in the wider context, and shows a degree of originality or professionalism. Typically this is work that would be considered of publishable standard.</td>
</tr>
<tr>
<td><strong>Bioinformatics strategy</strong></td>
<td>Strategy applied contains numerous errors and largely inappropriate</td>
<td>Analysis is basically correct. Some of the interpretation might be weak – some issues in developing a convincing and supportable hypothesis around the effect of the variation from the analyses undertaken.</td>
<td>A good analysis that covers the expected areas and provides a solid interpretation. Some evidence of a consideration of standards and governance rules applying to the analysis.</td>
<td>An excellent analysis that uses appropriate tools, develops a properly supported hypothesis, and shows a good understanding of the standards and governance requirements of a clinical bioinformatics pipeline.</td>
</tr>
<tr>
<td><strong>Medical consequences</strong></td>
<td>Little or no understanding of the consequences of the analysis for patient care</td>
<td>An attempt is made to determine the consequences, but poorly supported from other sources.</td>
<td>A good discussion of the consequences, with some support from other sources.</td>
<td>An excellent interpretation of the consequences, with good support from a range of external sources.</td>
</tr>
<tr>
<td><strong>Report quality</strong></td>
<td>Numerous errors with little referencing. Poorly structured or disorganised</td>
<td>Report has a good basic structure, minimal careless errors and appropriate presentation</td>
<td>A well written report that is well structured and with good referencing.</td>
<td>A report that is written and presented to a very high standard. Good use is made of figures and tables. Referencing complete and appropriate.</td>
</tr>
</tbody>
</table>