BioS

DIGITAL SKILLS ON COMPUTATIONAL BIOLOGY FOR HEALTH PROFESSIONALS

Project No: 591945-EPP-1-2017-1-DE-EPPKA2-SSA

R3.4. BioS trainee's handbook

WP3: Development of the BioS Curricula Syllabi and Educational Material (OER)

Responsible Partner: P5 BIB

Co-funded by the Erasmus+ Programme of the European Union

Co-funded by the
Erasmus+ Programme
of the European Union

Project information

Project acronym:	BioS
Project title:	Digital Skills on Computational Biology for Health Professionals
Agreement number:	2017-3424/591945-EPP-1-2017-1-DE-EPPKA2-SSA
EU programme:	Erasmus+ KA2-Sector Skills Alliances (SSA)
Project website:	bios-project.eu

Prepared by

Authoring Partner(s):	BIB

BioS
DIGITAL SKILLS ON COMPUTATIONAL BIOLOGY FOR HEALTH PROFESSIONALS

Table of contents

BioS
DIGITAL SKILLS ON COMPUTATIONAL BIOLOGY FOR HEALTH PROFESSIONALS

1. Introduction to BioS Courses

The BioS courses aim to provide physicians with knowledge, skills and competencies in the field of bioinformatics that will enable them to improve their clinical practice. Therefore, they include both the basic concepts of bioinformatics and computational biology, as well as the most recent advances in this field that can have a clear application for physicians and transform their practice. The courses have a modular and flexible structure. Four modules are offered, which can be taken sequentially or not, depending on the level of knowledge in bioinformatics of the participants. However, it is advisable to take module 1 (introduction to bioinformatics) to make the best use of the following three modules. The courses are based on an active and contextualised learning model, in which students are the centre of learning and learn through diverse authentic activities, close to their professional reality.

The teaching materials have been developed with a scientific team with expertise in the field of bioinformatics, the advice of an advisory committee formed with different stakeholders (different profiles of medical doctors, bioinformaticians, researchers and e-learning experts) and technical and pedagogical support.

Table 1. The teaching team of BioS modules

Modules	Coordinador	Teachers
Module 1	Dr.Cedric Notredam	Dr.Cedric Notredam Group leader of Notredam Lab- Comparative genomics at Centre for Genomic Regulation-CRG in Barcelona, Spain.
Module 2	Dr. Hafid Laayouni	Dr. Hafid Laayouni Associate teacher and researcher at Universitat Pompeu Fabra. Coordinator of the Bachelor degree of Bioinformatics in Barcelona, Spain. Dra. Laura Serra Associate teacher od Statistics at Universitat de Girona (UdG) in Girona, Spain. Dr. Oscar Lao Group leader of Population Genomics research at the CNAG-CRG in Barcelona, Spain.
Module 3	Dr. Ferran Casals	Dr. Ferran Casals Head of the Genomics Core Facility at Universitat

		Pompeu Fabra, associate teacher and researcher at the same University in Barcelona, Spain. Dr. Gerard Muntaner Postdoctoral scientist at the Universitat Rovira i Virgili (URV) in Tarragona, Spain Dra. Clara Serra Genetic counselor at the Vall d'Hebron University Hospital in Barcelona, Spain. Dra. Ivon Cusco Researcher at the genètic diagnosis laboratory at the Vall d'Hebron University Hospital in Barcelona, Spain.
Module 4	Dra. Clara Serra	Dr. Ferran Casals Dr. Gerard Muntaner Dra. Clara Serra Dra. Ivon Cusco

2. Description of the modules

The BioS curriculum is composed of the following modules: 1) Introduction to bioinformatics, 2) Computational statistics for clinical doctors, 3) Personalized genomics in patient care, and 4) Quality improvement in medical care.

The key learning objectives of the 4 modules and the dedication of workload expected for each module is described below. The knowledge, skills and competences to be developed in each module are specified in Annex 1.

Table 2. Modules Learning Outcomes and Workload

Modules	Learning outcomes (LO)	Weeks/ Workload
1. Introduction to bioinformatics	- Retrieve information and data regarding specific genes and proteins which could be chosen as candidate genes for a disease, e.g. functional information and sequence variant information - Perform analyses and comparisons to identify essential and non-essential parts in a gene or a protein, e.g. multiple sequence alignments using publicly available, web-based tools - Understand how applying such skills can lead to quick and cost-effective diagnoses of diseases and disorders with a genetic component	4 weeks 35-40h
2. Computational statistics for clinical doctors	- Understanding of descriptive statistics, elements of probability, hypothesis testing, nonparametric methods, correlation analysis, and linear regression - Understanding of how to choose appropriate statistical tests and how to assess statistical significance - Understanding of how to visualize data and carry out statistical testing - Understanding of R, a powerful programming language for statistical computing and graphics - Understanding selected kinds of analyses of biomedical data that a professional can carry out easily using e.g. the package of R for the analysis of sequencing data from a patient	4 weeks 35-40h
3. Personalized genomics in	- Explain single nucleotide polymorphisms (SNPs) and different types of SNPs - Describe some example roles of SNP's in disease	4 weeks 35-40h

patient care	development - Describe the basic principles of variant effect prediction for genetic variants in protein-coding genes and in micro-RNA genes - Use reports in patient care from commercial personal genomics services - Interpret SNP-related increased and decreased risks in selected diseases - Interpret commercial reports and translate them to actions in appropriate health care segments	
4. Quality improvement in medical care	- More efficient communication of disease risks related to genetic testing	2 weeks 5-10h

The modules are structured by weeks and contain the following activities:

- **Interviews with experts**: each module is introduced with an interview with a renowned researcher in the field. The objective of the interview is to provide an overview of the state of the art of the topic, its relevance and applicability for clinical practice.

- **Video lessons**: every week there are two lessons taught by of expert teachers, where they explain in a synthetic way the basic concepts that will be worked during the week. They last about 10 minutes each. It is important to watch the videos and understand them in order to be able to do the rest of the activities proposed for the week.

- **Short problems**: every week there is 1 short problem that has the purpose of applying the basic concepts that have been taught in the videos and consolidate them. Once you have submitted your answer, you will automatically receive the correct answer so that you can self-assess.

- **Learning activities**: every week there is a learning activity where the acquired knowledge through the video lessons must be applied and deepened. These are case studies, based on the real experiences of bioinformatics experts. These activities will take a little more time than the short problems. In some cases, you will have to read an article to be able to do the activity, in others you will have to analyze or interpret real biological data. By doing these activities you will be able to really develop the competences planned for the course and be able to apply

them further in your professional practice. Once you have submitted your answer, you will receive the correct answer so that you can self-evaluate

- **Readings of articles or reference materials**: each week there are proposed readings to go deeper into the knowledge of the topic. In some cases, it is necessary to read the documents in order to carry out the practical cases. When this is the case, it is already specified in the case studies. In other cases, literature is recommended for those students who want to go a little further. In all cases, it is worth consulting it

- **Quizzes**: at the end of each week there is a quiz of 10 questions to check if the basic knowledge has been acquired. Once you complete the quiz, you will receive the correct answers so that you can self-assess

- **Forum**: In each module there is at least one discussion activity, which aims to discuss real situations related to the application of bioinformatics in clinical practice. These activities will allow you to meet other course participants and exchange experiences and opinions. It is highly recommended that you participate in them, as you can learn a lot from your colleagues.

3. The training materials

3.1 Module 1: Introduction to bioinformatics

This module aims to introduce the basic knowledge of how molecular data available in modern biomedicine can be used to promote genomic/personalised medicine. The course introduces the student to understanding, managing and analyzing different types of data, from genomes, to DNA sequences and proteins or protein structures. The main biological databases are introduced and basic methods for their analysis are taught.

The module is introduced through an interview with Roderic Guigó, a reputable researcher in the field of bioinformatics. He is the group leader of Computational Biology of RNA Processing research group in Centre for Genomic Regulation (CRG) in Barcelona, Spain.

In the following table, the topics of the video classes contained in the module can be seen, as well as the lecturer who teaches them.

Table 3. Video lectures for Module 1. Introduction to bioinformatics

Week	Topic	Video	Title	Teacher
1	Genes within Genomes	111	Overview of bioinformatics with respect to molecular biology and genetics concepts	Cedric Notredam
		112	Navigating genomes	Cedric Notredam
2	Searching Databases for Functional Information	121	Gathering information in medical genetic databases	Cedric Notredam
		122	Gathering information in Generic Bioinformatics databases	Cedric Notredam
3	Exploring RNA and Protein products of a gene	131	Finding out about the level of expression of a gene	Cedric Notredam
		132	Finding out about the protein	Cedric Notredam
4	Healthy and Unhealthy proteins	141	Homology searches and multiple alignments in Biology	Cedric Notredam
		142	Metabolic pathways and integrated cellular function	Cedric Notredam

In order to acquire the skills of the module, the contents covered in the video-lessons must be put into practice. That is why different types of activities are proposed, from the simplest to the most complex. In the following table you can see all the activities proposed for module 1.

Table 4. Learning activities for Module 1. Introduction to bioinformatics

Activities	Week 1	Week 2	Week 3	Week 4
Short problem	1	1	1	1
Learning activity	1	1	1	1
Quiz	1	1	1	1
Readings	3			
Forum activity				1

3.2. Module 2: Computational Statistics for clinical doctors

The aim of this module is to provide a practical introduction to the analysis of Big data biomedical field, in order to acquire a critical understanding of the reliability of analysis results. In this module, medical doctors will be able to understand and become familiar with the statistical environment of R and apply it to the analysis of biological data in an efficient manner.

The module is introduced through an interview with researcher Antonio Barbadilla, who is the research leader of the Bioinformatics of Genomic Diversity group in the the Institute of Biotechnology and Biomedicine at the University Autònoma of Barcelona (Spain).

In the following table, you can see the topics of the video classes contained in the module, as well as the lecturers who teach them.

Table 5. Video lectures for Module 2. Computational Statistics for clinical doctors

Week	Topic	Video	Title	Teacher
1	Descriptive statistics	211	Overview of type of variables (qualitative and quantitative) and how to deal with them	Laura Serra
		212	Introduction to R	Laura Serra

2	Statistical inference	221	Overview of statistical significance	Oscar Lao
		222	Hypothesis testing in medical practice	Oscar Lao
3	Linear regression	231	Introduction to linear regression	Laura Serra
		232	Linear regression models using R	Laura Serra
4	Categorical Data	241	Introduction to Categorical Data analysis	Hafid Laayouni
		242	Summary on the statistical inference interpretation	Hafid Laayouni

In order to acquire the skills of the module, the contents covered in the video-lessons must be put into practice. That is why different types of activities are proposed, from the simplest to the most complex. In the following table you can see all the activities proposed for module 2.

Table 6. Learning activities for Module 2. Computational Statistics for clinical doctors

Activities	Week 1	Week 2	Week 3	Week 4
Short problem	1	2	1	1
Learning activity	1	2	1	1
Quiz	1	2	1	1
Readings	4	1	4	
Forum activity		1		

3.3. Module 3. Personalized genomics in patient care

The purpose of this module is to provide medical doctors the necessary knowledge and skills to interpret results from genetic analyses and commercial personalized genomics services, like 23andMe, deCODE, Gene by Gene, etc. This module facilitates integrating these services into their patient care activities.

The module is introduced through an interview with researcher Nuria Lopez-Bigas, which is the group leader of Biomedical Genomics Research Group in Institute for Research in Biomedicine in Barcelona (Spain).

In the following table, you can see the topics of the video classes contained in the module, as well as the lecturers who teach them.

Table 7. Video lectures of Module 3. Personalized genomics in patient care

Week	Topic	Video	Title	Teacher
1	Inheritance Model of Diseases	311	Analysis of pedigrees. Mendelian models	Clara Serra
		312	Other models of inheritance	Ivon Cusco
2	Human Genetics sources of variability	321	Genetic variation	Ferran Casals
		322	Population variation or Disease causing variants	Gerard Muntaner
3	Diagnostic tools: How to select the correct test. Cytogenetics	331	Cytogenetics and molecular cytogenetics	Clara Serra
		332	Cytogenetics and molecular cytogenetics. Data analysis and interpretation	Ivon Cusco
4	Diagnostic tools: How to select the correct test. Sequencing	341	Nucleotide variants	Ferran Casals
		342	Interpretation of genomic analysis	Gerard Muntaner

In order to acquire the skills of the module, the contents covered in the video-lessons must be put into practice. That is why different types of activities are proposed, from the simplest to the most complex. In the following table you can see all the activities proposed for module 3.

Table 8. Learning activities for Module 3. Personalized genomics in patient care

Activities	Week 1	Week 2	Week 3	Week 4
Short problem	1	1	1	1
Learning activity	1	1	1	1
Quiz	1	1	1	1
Readings	3			
Forum activity				1

3.4. Module 4. Quality improvement in Healthcare

This module programme will aim to equip trainees with a range of knowledge and skills, which are relevant and applicable in communications within healthcare contexts. Participants will learn how to build high-performing and engaged healthcare teams, establish and sustain effective clinical relationships, as well as implement strategies and tools to support patient-centered care. Additionally, with patient safety initiatives at the forefront of care, a major goal of this module will be to help health care professionals to develop the background knowledge and skills necessary for the specialty of risk management. This module is focused especially to the communication, ethics and risks associated with genetic testing and disease risk assessment.

The module is introduced through an interview with the researcher Jordi Surrelles, group leader of Genomic Instability and DNA Repair Research Group in Universitat Autònoma de Barcelona (Spain).

In the following table, you can see the topics of the video classes contained in the module, as well as the lecturers who teach them.

Table 9. Video lectures of Module 4. Quality improvement in Healtcare

Week	Topic	Video	Title	Teacher
1	Risk assessment	411	Genetic risk assessment. Bayesian analysis	Ferran Casals
		412	Genetic risk prediction in complex disorders	Gerard Muntaner
2	Ethical and Communication skills	421	Ethical and Communication skills	Clara Serra

In order to acquire the skills of the module, the contents covered in the video-lessons must be put into practice. That is why different types of activities are proposed, from the simplest to the most complex. In the following table you can see all the activities proposed for module 4.

Table 10. Learning activities for Module 4. Quality improvement in Healtcare

Activities	Week 1	Week 2
Short problem	1	
Learning activity	1	1
Quiz	1	1
Readings	5	
Forum activity		1

4. Grading policy

The final grade is calculated based on two categories:

1. The score obtained from the practical quizzes. There are 13 practical quizzes throughout the 4 modules.
2. The score obtained from the Final Exam quizzes. There are 4 Final Exam quizzes, one per each module.

Each category contributes to 50% in the final grade. The passing threshold of the course is set to 50% of the final grade.

Appendix 1. Learning outcomes for the four modules

Table 1: Module 1: Learning Outcomes of Introduction to Bioinformatics

5	INTRODUCTION TO BIOINFORMATICS	40	4
KEY TECHNICAL OUTCOME	**KNOWLEDGE**	**SKILLS**	**COMPETENCE**
1. Retrieve information and data regarding specific genes and proteins which could be chosen as candidate genes for a disease.	The Learner knows and understands: ● The relevance of biological sequences for health and diseases ● Basic terms related to sequence handling ● Medical relevance of sequence annotation	The Learner is able to: ● Gather information on selected genes and proteins using tools such as BLAST, UniProt, and PDB ● Compare biological sequences through multiple sequence alignment ● Identify the active site of an HIV Serine protease (or of any other structurally characterized enzyme)	The Learner: ● Is aware of the power of modern high-throughput sequencing methods and applies sequencing data to improve diagnostics of diseases with a suspected genetic component ● Uses up-to-date knowledge from databases of genes and proteins to support their estimates of the significance of genes suggested as contributors in a genetic disease ● When needed, considers protein structure in the interpretation of variant data
	PERFORMANCE CRITERIA		
	● Uses sequence and annotation files to access up-to-date, in-depth knowledge of medically relevant genes (with the help of the demonstrated Internet databases and tools) ● Uses the discovered information to make more informed clinical decisions based on genetic variant data		
KEY TECHNICAL OUTCOME	**KNOWLEDGE**	**SKILLS**	**COMPETENCE**
2. Visualize genomic features and perform simple analysis on them using Internet based tools.	The Learner knows and understands: ● List of available databases and other Internet resources ● Functional Genome Annotation and Metabolic Pathways	The Learner is able to: ● Visualize genomic data in a genome browser ● Find information of known genome variants associated to desired genes ● Find and visualize functional genome annotations and metabolic pathway information	The Learner: ● Uses genomic tools routinely to get a quick, integrated view of data related to any gene as necessary ● Gains a knowledge-based, data driven view of every new disease-related gene they encounter
	PERFORMANCE CRITERIA		
	● Uses the demonstrated tools to visualize annotations of their candidate gene lists; ● Uses Internet resources including Genome Browsers and KEGG; ● Interprets gene information critically in the light of up-to-date information.		
OUTPUTS			
• **More confidence in data from modern-day high-throughput sequencing;** • **Making more use of such data in clinical decision-making;** • **More informed interpretation of personal genome data;** • **Quicker and more relevant diagnoses of diseases with an assumed genetic component.**			

Table 1: Module 2: Learning Outcomes of Computational Statistics for Clinical Doctors

EQF LEVEL	UNIT OF LEARNING OUTCOME	HOURS	CREDITS
5	COMPUTATIONAL STATISTICS FOR CLINICAL DOCTORS	40	4

KEY TECHNICAL OUTCOME	KNOWLEDGE	SKILLS	COMPETENCE
1. Understanding key elements of modern statistical analysis.	The Learner knows and understands: ● Descriptive statistics, elements of probability, hypothesis testing, nonparametric methods, correlation analysis, and linear regression ● Elements of statistical reports ● Elements of the visualization of statistical data	The Learner is able to: ● Choose appropriate statistical tests ● Perform simple statistical analyses using software tools ● Assess statistical significance ● Evaluate if appropriate statistical test is used in an analysis ● Interpret statistical graphs	The Learner: ● Pays attention to statistical values associated with reports of sequencing data and its analysis reports; ● Weights their decision-making in the light of the statistical uncertainty of each finding.
	PERFORMANCE CRITERIA		
	● Operates with elements of statistical analysis to interpret analysis results; ● Reads analysis reports and research articles with a statistically critical approach.		

KEY TECHNICAL OUTCOME	KNOWLEDGE	SKILLS	COMPETENCE
2. Using free software for statistical analysis of data from case studies.	The Learner knows and understands: ● The role of R software in statistical analysis ● CRAN and Bioconductor	The Learner is able to: ● Install R, RStudio, and selected related statistical software packages and use them ● Organize and perform a data analysis project of biomedical data ● Create and handle graphs from an analysis	The Learner: ● Is aware of the power and limitations of statistical analyses of biomedical data, especially of high-throughput sequencing data ● Requests custom-made analyses from professionals, knowing that efficient, simple, and free tools exist enabling almost any analysis they can think of, in an easy and cost-efficient way
	PERFORMANCE CRITERIA		
	● Makes more use of statistical parameters in decision-making, especially related to analysis of sequencing data from a patient ● Designs ideas for custom analyses if needed in complex situations		

OUTPUTS
• **More efficient use of biomedical research results;** ● **Incorporating sequencing data and the use of associated statistical data more efficiently into diagnostic practice.**

BioS
DIGITAL SKILLS ON COMPUTATIONAL BIOLOGY FOR HEALTH PROFESSIONALS

Table 2: Module 3: Learning Outcomes of Commercial personalized genomics services in patient care

EQF LEVEL	UNIT OF LEARNING OUTCOME	HOURS	CREDITS
5	COMMERCIAL PERSONALIZED GENOMICS SERVICES IN PATIENT CARE	40	4

KEY TECHNICAL OUTCOME	KNOWLEDGE	SKILLS	COMPETENCE
1. Understanding the nature and role of single nucleotide polymorphisms (SNPs) and other genetic variants.	The Learner knows and understands: ● Different levels of genetic and genomic variants; ● Variant terminology; ● Theoretical aspects of human genetics related to genomic variations; ● Variant analyses within populations; ● Diagnostic tools used for variant detection and analysis.	The Learner is able to: ● Access and study genetic variant data from on-line databases; ● Read reports of patient sequence variant analyses.	The Learner: ● Considers carefully the contribution of genetic variants for a given patient case
	PERFORMANCE CRITERIA		
	● Uses this learning for more informed reading of reports from genomic sequencing services ● Uses variant databases for accessing more information on disease gene candidates		

KEY TECHNICAL OUTCOME	KNOWLEDGE	SKILLS	COMPETENCE
2. Using variant analyses	The Learner knows and understands: ● Types of available commercial personalized genomics services and other genetic tests ● Differences between their analysis methodology ● Principles of variant effect predictions	The Learner is able to: ● Select appropriate genetic tests for a given clinical situation ● Interpret reports from different genomics services ● Assess the reliability of information sources used in different genomic services	The Learner: ● Evaluates always if genetic variant analyses are needed and appropriate; ● Uses personal genomic reports to gain a better understanding of patient's health status; ● Applies available research knowledge for personalizing patient treatment and/or preventive measures.
	PERFORMANCE CRITERIA		
	● Uses different genomics services to support clinical work when a genetic component is assumed ● Assesses individual patient status taking into account information from genomic sequencing reports		

OUTPUTS

- **Integration of the available commercial personalized genomics services into patient care practice**
- **More timely and more cost-efficient clinical decisions**
- **Better choices in treatments and in disease prevention**
- **Improved health**

Table 3: Module 4: Learning Outcomes of Quality Improvement in Healthcare

EQF LEVEL	UNIT OF LEARNING OUTCOME	HOURS	CREDITS
5	QUALITY IMPROVEMENT IN HEALTHCARE	10	2

KEY TECHNICAL OUTCOME	KNOWLEDGE	SKILLS	COMPETENCE
More efficient communication of disease risks related to genetic testing	The Learner knows and understands: ● Concepts related to disease risk assessment ● Ethical issues related to genetic testing ● Tools to support communication	The Learner is able to: ● Communicate risk information to patients in understandable and compassionate ways ● Help health professionals make informed decisions of treatment or non-treatment of discovered diseases ● Give genetic counselling	The Learner: ● Identifies and addresses the differences in patients' values, preferences and expressed needs ● Aims at a coaching culture that supports consistent exceptional care and service

PERFORMANCE CRITERIA
● Implements strategies and tools to support patient-centred care; ● Improves the patient experience by better communication.

OUTPUTS
• **Attracting and engaging customer-focused employees who are passionate about providing the best and most compassionate, yet efficient, care to the patient**
• **Effective clinical relationships**
• **Patient experience framework that better meets and exceeds the patient's need**