Course Title: Statistical data analysis and research methods
Course Subject Code: BMI **Course Number:** 504
Type of Instruction: SEM **Class Number:** 20048
Semester: Spring 2019

DRAFT!!! To be adjusted after 1st class

1. *Course Information*

- Date(s)/Time(s): Thursdays 10AM – 12.30/1PM
- Delivery Mode: Traditional
- Number of Credits: 3
- Instructors

Course director: Werner Ceusters, MD (contact: 77 Goodell street, 5th floor, on appointment: wceusters@gmail.com)

Lecturers:

Statistics: Sarah Mullin, MSc (contact: 77 Goodell street, 5th floor, on appointment: sarahmul@buffalo.edu)

All other topics: Werner Ceusters, MD (contact: 77 Goodell street, 5th floor, on appointment only through wceusters@gmail.com)

2. *Course Description*

- This course provides an introduction to research design and methods in biomedical informatics and aims to enhance the students' quantitative and qualitative research skills. Six themes will be covered: (1) the fundamentals of scientific research; (2) elements of philosophy of science, (3) quantitative research methods including statistics; clinical epidemiology, population studies and big data; (4) theoretically informed qualitative research; (5) integration of research methods in biomedical informatics; and (6) research ethics, including informed consent, and role of IRBs.
- Each class in the course consists of a theoretical and practical part, either in-class, or in the form of an assignment. The theoretical part will be in the form of interactive lectures surveying the formulation of research questions, the development of testable hypotheses, the selection and application of appropriate research designs and methods, data collection and analysis methods. These skills can then be applied in the practical part in which students will develop a research proposal for a topic in their interest.
- At the end of the course, students should be able to apply the methods taught in subsequent courses and research projects and use them for the evaluation and production of research proposals and papers.
- Course prerequisites: none.

3. *Student Learning Outcomes (SLO)*

Course Learning Outcomes; students will be able to:	**BMI Graduate Program Outcomes / Competencies**	**Instructional Method(s)**	**Assessment Method(s)**
1. Explain the fundamentals of scientific inquiry	O1:The most widely used clinical and informatics research methods	Lecture: C1, C2	In-class test: T1
2. Articulate research questions	O2:The generation, acquisition, modeling, representation, and management of evidence-based knowledge sources for decision support O3:The characteristics of public health data as distinguished from clinical healthcare data	Lecture: C4, C6	Assignment: A2, A3
3. Assess the quality of quantitative and qualitative studies	O1:The most widely used clinical and informatics research methods O4:Methods of data representation, manipulation, storage, analysis and	Lecture: C3, C8	In-class test: T2, T4 Assignment: A1

	mining in healthcare and biomedical research databases		
4. Understand and apply the concept of hypothesis in quantitative studies	O1:The most widely used clinical and informatics research methods. O5:Technical approaches to acquiring, modeling, representing and managing healthcare and biomedical research knowledge	Lecture: C4, C10	Assignment: A2
5. Explain the role of causal theories in the design and interpretation of quantitative studies	O6:Research & data management methods with large clinical populations, including clinical trials O7:The essential components of clinical and biomedical data statistical analysis	Lecture: C7, C10	Assignment: A4
6. Understand and use statistical methods for calculating summary estimates, measures of variability and confidence intervals	O7:The essential components of clinical and biomedical data statistical analysis	Lecture: C6, C8, C9	Assignment: A3, A5 In-class test: T4
7. Understand probabilities and discrete and continuous distributions	O7:The essential components of clinical and biomedical data statistical analysis	Lecture: C9	Assignment: A5
8. Carry out and interpret a variety of tests of significance	O7:The essential components of clinical and biomedical data statistical analysis	Lecture: C9	Assignment: A5
9. Understand and use power and sample size calculations	O7:The essential components of clinical and biomedical data statistical analysis	Lecture: C9	Assignment: A5
10. Use theory in qualitative research	O1:The most widely used clinical and informatics research methods	Lecture C5	Open-book test: T3
11. Explain different forms of qualitative inquiry, including interviews, focus groups and observations and understand their benefits and limitations	O8:Information retrieval and critical analysis skills	Lecture C5, C7, C11	Open-book test: T3 Assignment: A4, A6
12. Explain and use different tools for qualitative data analysis	O7:The essential components of clinical and biomedical data statistical analysis	Lecture: C11	Assignment: A6
13. Judge the suitability of different statistical methods to describe research findings	O7:The essential components of clinical and biomedical data statistical analysis	Lecture: C10, C11	Assignment: A6
14. Construct a coherent research proposal that includes an abstract, introductions, literature review, research questions, ethical considerations, and methodology	O9:Ethical theories and challenges in Biomedical Informatics	Lecture: C4, C10, C13	In-class exercise: C12 In-class test: T5 Assignment: A7 Student presentation

4. COURSE REQUIREMENTS

- Students are required to read 12 papers/book chapters/publications as listed in the course materials below.
- Some classes will start with a test to assess the student's preparedness for the class, and a post-test to assess his attention to the lecture. Students absent for these classes will receive a 0% score for these tests unless the instructor and the course director have been informed through email of a valid reason for absence prior to the beginning of the class. When notified of absence in due time, the instructor may propose an alternative test or assignment for that class.
- For several lectures, availability of laptop (hands-on, open-book tests) is required.
- All assignments need to be completed prior to the deadline specified in the course schedule and send electronically to wceusters@gmail.com. Assignments ordered by Mullin should, in addition to be sent to wceusters@gmail.com, also be sent to sarahmul@buffalo.edu. The 'sent' date in the metadata header of the message through which the completed assignments will be received will be taken for assessment of in-time delivery. Whenever an assignment is delivered past the due date, a

penalty of 1% of the positive final score will be applied. This penalty will not be applied if the assignment is not delivered at all. Students unable to meet the deadline must inform the course director and instructor by email about a valid reason prior to reaching the deadline. An alternative assignment and/or due date can then be agreed upon.

5. GRADING POLICY

- Grading follows standard graduate policies (http://grad.buffalo.edu/Academics/Policies-Procedures/Grading-Procedures.html)
- Learning assessments will be graded based on rubric criteria and weighted according to the following detailed break-down. If the final results for all students are outside the expected range, curve grading might be used at the discretion of the course director.

Class	Date	Instructor	Assessment	Due dates	Score weight
C1	31-Jan	Ceusters (1)			
C2	07-Feb	Ceusters (2)	in-class test T1		2%
C3	14-Feb	Ceusters (3)	in-class test T2 assignment: A1	 Feb 19 - noon	2% 3%
C4	21-Feb	Ceusters (4)	assignment: A2	Feb 26 - noon	3%
C5	28-Feb	Ceusters (5)	test at home: T3	Mar 5 - noon	2%
C6	07-Mar	Ceusters (6)	assignment: A3	Mar 12 - noon	2%
C7	14-Mar	Ceusters (7)	assignment: A4	Mar 26 - noon	5%
C8	28-Mar	Ceusters (8)	in-class test T4		2%
C9	04-Apr	Mullin (1)	assignment: A5	Apr 09 - noon	3%
C10	11-Apr	Ceusters (9)			
C11	18-Apr	Ceusters (10)	assignment: A6	Apr 24 - noon	3%
C12	25-Apr	Ceusters (11)	In-class exercise part 1 In-class exercise part 2		3% 2%
C13	02-May	Ceusters (12)	in-class test T5 assignment: A7		3% 20%
C14	09-May	Ceusters (13)	Final presentation		20%
EXAM	16-May	Ceusters (14)			25%
				Total	**100%**

Final Grades:

Grade	Quality Points	Percentage
A	4.0	93.0% -100.00%
A-	3.67	90.0% - 92.9%
B+	3.33	87.0% - 89.9%
B	3.00	83.0% - 86.9%
B-	2.67	80.0% - 82.9%
C+	2.33	77.0% - 79.9%
C	2.00	73.0% - 76.9%
C-	1.67	70.0% - 72.9%
D+	1.33	67.0% - 69.9%
D	1.00	60.0% - 66.9%
F	0	59.9 or below

An interim grade of Incomplete (I) may be assigned if the student has not completed all requirements for the course. An interim grade of 'I' shall not be assigned to a student who did not attend the course. The default grade accompanying an interim grade of 'I' shall be 'U' and will be displayed on the UB record as 'IU.' The default Unsatisfactory (U) grade shall become the permanent course grade of record if the 'IU' is not changed through formal notice by the instructor upon the student's completion of the course.
Assignment of an interim 'IU' is at the discretion of the instructor. A grade of 'IU' can be assigned only if successful completion of unfulfilled course requirements can result in a final grade better than the default 'U' grade. The student should have a passing average in the requirements already completed. The instructor shall provide the student specification, in writing, of the requirements to be fulfilled.

6. *ACADEMIC INTEGRITY*

- Academic integrity is a fundamental university value. Through the honest completion of academic work, students sustain the integrity of the university while facilitating the university's imperative for the transmission of knowledge and culture based upon the generation of new and innovative ideas. See http://grad.buffalo.edu/Academics/Policies-Procedures/Academic-Integrity.html.
- Students may collaborate for the assignments in which case the submitted materials should be clearly labeled as such, with the names of all collaborating students. In case students who collaborate cannot come to a consensus for certain parts of the work, alternate solutions proposed by individual students should be clearly marked as such. Grading of individual students will take into account such alternatives.

7. *ACCESSIBILITY RESOURCES*

If you have any disability which requires reasonable accommodations to enable you to participate in this course, please contact the Office of Accessibility Resources, 25 Capen Hall, 645-2608, and also the instructor of this course .. The office will provide you with information and review appropriate arrangements for reasonable accommodations. http://www.student-affairs.buffalo.edu/ods/

8. *COURSE FEES*

Standard UB tuition and fees. No extra costs, except for students wishing to purchase the statistics handbook.

9. *Course Organization / Schedule*

Reference: http://registrar.buffalo.edu/calendars/academic/

C1. Jan 31 / Ceusters / Course introduction - Fundamentals of science and research / SLO 1

Class structure:

a) Participant and instructor introduction
b) Course introduction, housekeeping rules, expectations, course project work
c) Traditional lecture on (1) philosophy of science and research and (2) the scientific method

Post-class assignment:

Required reading:

R1. M Delgado-Rodríguez and J Llorca.
Bias.
J Epidemiol Community Health 2004;58:635-641
http://jech.bmj.com/content/58/8/635.full.pdf+html
!!! This paper will be the topic of a closed book in-class test during class C2 !!!

C2. Feb 7 / Ceusters / Types of Bias / SLO 1

Class structure:

a) T1: Assessment of R1: in-class test (closed book)
b) Interactive lecture on various types of bias

Post-class assignment:

Required reading:

R2. John P. A. Ioannidis
Why Most Published Research Findings Are False
PLOS Medicine 2005;2(8):e124
http://robotics.cs.tamu.edu/RSS2015NegativeResults/pmed.0020124.pdf
!!! This paper will be the topic of a closed book in-class test during class C3 !!!

R3. University of Southern California.
Organizing Your Social Sciences Research Paper: Types of Research Designs.
http://libguides.usc.edu/c.php?g=235034&p=1559832

C3. Feb 14 / Ceusters / Types of research designs / SLO 3

Class structure:

a) T2: Assessment of R2: in-class test (closed book) followed by discussion of correct answers.
b) Lecture introducing various research designs
c) Organization of assignment

Post-class assignment:

a) A1: Each student is assigned 3 research designs discussed in class and in paper R1. The task is to build a 4-column matrix with the 1st column containing the design steps and features concerning at least 1 of the 3 designs, and the remaining columns, one for each research design, an annotation regarding (1) whether, and if so, how, this step/feature figures in the design, (2) the extent to which it is impacted by any of the 6 corollaries discussed in paper R1, and (3) what should be done to minimize the effect.
Due date: Feb 19 – noon.
b) Required reading:
R4. Reinhold Haux *et.al.*
Research Strategies for Biomedical and Health Informatics: Some Thought-provoking and Critical Proposals to Encourage Scientific Debate on the Nature of Good Research in Medical Informatics.
Methods Inf Med. 2017; 56(Open): e1–e10. Published online 2017 Jan 25.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5388922/pdf/ME-56-Open-16010125.pdf
R5. Hilde Tobi & Jarl K. Kampen.
Research design: the methodology for interdisciplinary research framework
Qual Quant (2018) 52:1209–1225
https://link.springer.com/content/pdf/10.1007%2Fs11135-017-0513-8.pdf
c) Prepare for discussion during class C4 topics for a course research project you are interested in.

C4. Feb 21 / Ceusters / Planning of research projects / SLO 2, 4, 14

Class structure:

a) Lecture: Framework for team science research projects
b) Initial discussion about possible research projects
c) Rough definition of global research project and individual contributions

Post-class assignment:

a) Required reading:
R6. John P. A. Ioannidis
Why Most Clinical Research Is Not Useful
PLoS Med 13(6): e1002049
http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002049
b) A2: After reading R6, and taking into account the content of all previous BMI504 lectures and papers, students must write an outline for their individual research project in the spirit of the scientific method thereby adapting the original ideas discussed in class in such a way that Ioannidis' features for assessing whether *clinical* research is useful come out positive for what they propose. This should be backed up by an initial literature study.
Due date: Feb 26 – noon.

C5. Feb 28 / Ceusters / Qualitative research methods: theory and data collection methods / SLO 10, 11

Class structure:

Lecture on common qualitative data collection methods (Document Review, Observation, Interview (face-to-face), Focus Group Discussion, Ethnography,…)

Post-class assignment:

a) T3: Open book test to do at home with questions partly covered in the class but for which complete answers can be found in the literature cited. Students must answer the questions and document their answers with quotes and correct citation.
Due date: Mar 5 – noon.
b) Required reading:
R7. Centers for Disease Control
Principles of Epidemiology in Public Health Practice, Third Edition. An Introduction to Applied Epidemiology and Biostatistics. Lesson 3: Measures of Risk. Sections 1 through 6.
https://www.cdc.gov/ophss/csels/dsepd/ss1978/lesson3/index.html

C6. Mar 7 / Ceusters / Elements of epidemiology / SLO 2, 6

Class structure:

Lecture covering essential notions in population studies such as incidence, prevalence, mortality ratios, validity, reliability, sensitivity, and specificity, etc…

Post-class assignment:

A3: Write a short essay about the inaccuracies that might arise in incidence and prevalence estimations on the basis of diagnostic codes retrieved from electronic healthcare records. Length doesn't matter, correct identification of issues and argumentation does!

Due date: Mar 12 – noon.

C7. Mar 14 / Ceusters / Introduction to data analysis of quantitative and qualitative variables / SLO 5, 11

Class structure: interactive lecture

Post-class assignment:

a) A4: Based on the slides and references used in lecture C7, students must reformulate their individual research project so that it satisfies all requirements with respect to (1) the precise research question, (2) the null-hypothesis, (3) the statistical null-hypothesis, (4) determination of all relevant variables and (5) theoretical and operational linkage.
Due date: March 26 – noon.

b) Required reading:

R8. Anthony McCluskey and Abdul Ghaaliq Lalkhen
Statistics II: Central tendency and spread of data
Continuing Education in Anaesthesia, Critical Care & Pain 2007;7(4):127-130.
http://ceaccp.oxfordjournals.org/content/7/4/127.full.pdf+html
!!! This paper will be the topic of a closed book in-class test during class C8 !!!

Mar 22: NO CLASS: SPRING RECESS

C8. Mar 28 / Ceusters / Descriptive and elementary statistics / SLO 3, 6

Class structure:

a) T4: Closed book test on paper R8
b) lecture covering the theories and applications of average value, median, mode, variance, standard deviation, inter-quartile range, skewness, kurtosis, histogram, box and whisker plot.

Post-class assignment:

Required readings:

R9. Philip R.O. Payne, Elmer V. Bernstam, and Justin B. Starren.
Biomedical informatics meets data science: current state and future directions for interaction.
JAMIA Open, 1(2), 2018, 136–141.
https://academic.oup.com/jamiaopen/article-pdf/1/2/136/26294661/ooy032.pdf

R10. Streiner, David and Norman, Geoffrey
Biostatistics: Bare Essentials
People's Medical Publishing House
Chapters 4, 7 and 10 (If you need probability review, refer to chapters 5 and 6.)
Accessible through UB Libraries login: http://ebookcentral.proquest.com/lib/buffalo/detail.action?docID=3386956

C9. Apr 4/ Mullin / Statistical analysis / SLO 6, 7, 8, 9

Class structure:

Lecture plus guided group work on Discrete and Continuous Distributions, the Normal Distribution, the Central Limit Theorem, Population vs Sample, Hypothesis Testing and Confidence Intervals (Z-test, t-test, Chi-Square Test, Fisher Exact, non-parametric), Power and Sample Size Calculations for One and Two-Sample Hypothesis tests.

Post-lecture assignment:

A5: For each student, a task will be designed addressing statistical problems relevant to the student's individual research project Due date: April 9 – noon.

C10. Apr 11 / Ceusters / Clinical trial design / SLO 4, 5, 13, 14

Class structure:

Lecture on cohort study design, clinical study design, analysis of clinical trials, randomized controlled clinical trials, sample size and power, survival analysis, missing data, Cox proportional hazard model, hazard ratio, kaplan meier, 2x2 factorial designs, cross over designs.

Post-lecture assignment:

Required reading:

R11. Lawrence A. Palinkas, Gregory A. Aarons, Sarah Horwitz, Patricia Chamberlain, Michael Hurlburt, John Landsverk Mixed Method Designs in Implementation Research
Adm Policy Ment Health (2011) 38:44–53
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3025112/pdf/10488_2010_Article_314.pdf

C11. Apr 18 / Ceusters /Mixed methods: Integration of quantitative and qualitative methods / SLO 11, 12, 13

Class structure:

a) The first part will be an interactive lecture covering the topic.
b) The second part will consist of a guided discussion aimed at determining the best research design to (dis)confirm hypotheses proposed for the individual research projects and to include an experimental design requiring a mixed design of not present in the original research proposal.

Post-class assignment:

a) Required reading:
   R12. The logic and structure of research proposals
   https://www.mheducation.co.uk/openup/chapters/9780335244065.pdf
a) A6: Students will update their proposal with an experimental design by building further on the guided discussion. Due date: April 24 – 5PM

C12. April 25 / Ceusters /Fundamentals of research: quality of research proposals / SLO14

Class structure: in-class exercise

a) during the first part, the class will use R12 and materials from all previous classes to develop an evaluation template for research proposals
b) during the second part, students will evaluate using the template the research proposals of their peers.

C13. May 2 / Ceusters / Research ethics; plagiarism; informed consent; IRB / SLO 14

Class structure: interactive lecture followed by in-class open book test (T5) on the topics covered in the lecture.

Post-class assignment:

Students will complete their final research proposal (A7) and prepare a powerpoint presentation of this proposal Due date: May 9 – 10AM

C14. May 9 / Ceusters / Presentation of final research proposals / SLO 14

Class structure: presentation of research proposals. Each student has 150 minutes divided by number of students participating, 75% of time for presentation, 25% for questions.

FINAL EXAM: May 16 / 9AM

The final exam will be held in the class room. It will be composed of questions and exercises covering the complete content of the course. Students may bring their laptop to the exam pre-loaded with any documentation they consider useful to consult during the text, whether or not used during the course. However, wifi services must be disabled and use of cell phones is not allowed, this to ensure that students will do the test individually.

10. COURSE MATERIALS

- This course requires the following papers and electronic publications, all of which are available publicly or through the UB Libraries:

R1. M Delgado-Rodríguez and J Llorca.
Bias.
J Epidemiol Community Health 2004;58:635-641
http://jech.bmj.com/content/58/8/635.full.pdf+html

R2. John P. A. Ioannidis
Why Most Published Research Findings Are False
PLOS Medicine 2005;2(8):e124
http://robotics.cs.tamu.edu/RSS2015NegativeResults/pmed.0020124.pdf

R3. University of Southern California.
Organizing Your Social Sciences Research Paper: Types of Research Designs.
http://libguides.usc.edu/c.php?g=235034&p=1559832

R4. Reinhold Haux *et.al.*
Research Strategies for Biomedical and Health Informatics: Some Thought-provoking and Critical Proposals to Encourage Scientific Debate on the Nature of Good Research in Medical Informatics.
Methods Inf Med. 2017; 56(Open): e1–e10. Published online 2017 Jan 25.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5388922/pdf/ME-56-Open-16010125.pdf

R5. Hilde Tobi & Jarl K. Kampen.
Research design: the methodology for interdisciplinary research framework
Qual Quant (2018) 52:1209–1225
https://link.springer.com/content/pdf/10.1007%2Fs11135-017-0513-8.pdf

R6. John P. A. Ioannidis
Why Most Clinical Research Is Not Useful
PLoS Med 13(6): e1002049
http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002049

R7. Centers for Disease Control
Principles of Epidemiology in Public Health Practice, Third Edition. An Introduction to Applied Epidemiology and Biostatistics. Lesson 3: Measures of Risk. Sections 1 through 6.
https://www.cdc.gov/ophss/csels/dsepd/ss1978/lesson3/index.html

R8. Anthony McCluskey and Abdul Ghaaliq Lalkhen
Statistics II: Central tendency and spread of data
Continuing Education in Anaesthesia, Critical Care & Pain 2007;7(4):127-130.
http://ceaccp.oxfordjournals.org/content/7/4/127.full.pdf+html

R9. Philip R.O. Payne, Elmer V. Bernstam, and Justin B. Starren.
Biomedical informatics meets data science: current state and future directions for interaction.
JAMIA Open, 1(2), 2018, 136–141.
https://academic.oup.com/jamiaopen/article-pdf/1/2/136/26294661/ooy032.pdf

R11. Lawrence A. Palinkas, Gregory A. Aarons, Sarah Horwitz, Patricia Chamberlain, Michael Hurlburt, John Landsverk
Mixed Method Designs in Implementation Research
Adm Policy Ment Health (2011) 38:44–53
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3025112/pdf/10488_2010_Article_314.pdf

R12. The logic and structure of research proposals
https://www.mheducation.co.uk/openup/chapters/9780335244065.pdf

- This course requires in addition the following book: Streiner, David and Norman, Geoffrey. Biostatistics: Bare Essentials. People's Medical Publishing House. Limited copies are available in the UB Library. Older versions might be found electronically on the web. We recommend purchasing.

R10. Streiner, David and Norman, Geoffrey
Biostatistics: Bare Essentials
People's Medical Publishing House
Chapters 4, 7 and 10 (If you need probability review, refer to chapters 5 and 6.)
Accessible through UB Libraries login: http://ebookcentral.proquest.com/lib/buffalo/detail.action?docID=3386956

11. ATTENDANCE POLICY

Students are expected to attend ***all*** lectures and lab exercises. For religious observances, university sanctioned events, athletic commitments and family/work obligations/emergencies, absences may be granted upon request but can have an effect on the finally obtained grade (see grading policy) unless additional coursework in replacement for missed graded activities is performed.
For course cancellation/emergency planning, see the university website for cancellations/delays due to weather or other unforeseen events (http://emergency.buffalo.edu/campus-weather-alerts.html)

12. CLASSROOM DECORUM

Students are expected to arrive in due time for each class. Most lectures will start with a pre-lecture test to assess the student's level of preparation for the class. This test contributes to the final grading. Use of cell phones and laptops is allowed for the purposes of the class, but not for private reasons. Additional rules of conduct, when applicable, will be explained by the instructors prior to the class.

13. UNIVERSITY SUPPORT SERVICES

Students are often unaware of university support services. For example, the Center for Excellence in Writing provides support for written work, and several tutoring centers on campus provide academic success support and resources.