A Course in Biology and Communication Skills for Master of Biostatistics Students

Jesse D. Troy[1,*], Josh Granek[1], Gregory P. Samsa[1], Gina-Maria Pomann[1], Sharon Updike[1], Steven C. Grambow[1] & Megan L. Neely[1]

[1]Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC, USA

*Correspondence: Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC, USA. Tel: 1-919-668-2932. E-mail: jesse.troy@duke.edu

Received: December 29, 2021 Accepted: March 11, 2022 Online Published: April 21, 2022

doi:10.5430/jct.v11n4p120 URL: https://doi.org/10.5430/jct.v11n4p120

Abstract

We describe an innovative, semester-long course in biology and communication skills for master's degree students in biostatistics. The primary goal of the course is to make the connection between biological science and statistics more explicit. The secondary goals are to teach oral and written communication skills in an appropriate context for applied biostatisticians, and to teach a structured approach to thinking that enables students to become lifelong learners in biology, study design, and the application of statistics to biomedical research. Critical evaluation of medical literature is the method used to teach biology and communication. Exercises are constructivist in nature, designed to be hands-on and encourage reflection through writing and oral communication. A single disease area (cancer) provides a motivating example to: 1) introduce students to the most commonly used study designs in medical and public health research, 2) illustrate how study design is used to address questions about human biology and disease, 3) teach basic biological concepts necessary for a successful career in biostatistics, and 4) train students to read and critically evaluate publications in peer-reviewed journals. We describe the design and features of the course, the intended audience, and provide detailed examples for instructors interested in designing similar courses.

Keywords: collaborative biostatistics, curriculum design, statistics education, biology education

1. Introduction

The successful practice of biostatistics requires a unique intersection of skills, including analytical skills, biological knowledge, and effective oral and written communication skills (Samsa et al., 2018). The primary focus of most graduate level training programs in biostatistics is on the analytical content, and with good reason. Indeed, biostatisticians cannot be effective without a well-developed analytical skillset. On the other hand, biostatisticians who have successfully integrated with multidisciplinary scientific teams realize that having biological knowledge and being able to effectively communicate or translate between statisticians and non-statistician scientists are also pre-requisites for effective collaboration (Zapf et al., 2019). In fact, one of the primary competencies required for collaborative biostatisticians is defined to be "Clinical and Domain Knowledge", a broad understanding of specific and relevant biomedical areas (Pomann et al., 2020). Given the importance of these skills to the success of future collaborative biostatisticians, university faculty who design biostatistics curricula often debate two key questions: 1) how much biology should we teach our students, and 2) how should we help our students communicate better?

Through our experience teaching biostatistics to master's degree students, the majority of whom join the workforce after graduation, we have come to a different point of view regarding these prototypical questions. Recently, we posited a model of biostatistical practice that conceptualizes the biostatistician as a life-long learner who is equipped not with some critical mass of encyclopedic knowledge, but who is an expert at connecting their domain-specific knowledge to new problems in the biological sciences through the establishment of a mental map; i.e., an approach to thinking (Troy et al., 2021). Our philosophy, therefore, is not to teach students "all they need to know" but rather to teach them how to gain knowledge efficiently and effectively in the future.

Toward this end, we designed a semester long course that teaches biological concepts and communication skills to master's degree students in biostatistics, loosely following the principles of constructivism (Biggs, 1996). The course

teaches how biostatisticians approach problems and navigate through the collaborative process (Samsa, 2018). The course is designed to use cancer as the primary application but employs methods that teach skills so they can be applied to any problem in the biomedical sciences. Importantly, the course uses critical evaluation of medical literature as a method for teaching biology and communication. This method is consistent with constructivist principles in that it follows a "see-one and do-one" approach and encourages reflection. The reflection is purposefully designed to be both written and oral to provide practice at communication, contextualized for the biomedical sciences.

This article describes the goals of the course and the student audience, gives an overview of the course design, and reviews the prominent features of the course. Specific examples are given in the appendix for instructors who are interested in developing similar courses. Finally, we share student feedback from end-of-semester surveys and discuss opportunities for expanding the course in the future.

2. Course Overview

2.1 Course Goals

Our primary goal for this course was to make the connection between biological science and statistics more explicit. Our secondary goals were to teach oral and written communication skills in an appropriate context for students who would be future applied biostatisticians, and to teach students a structured approach to thinking that enables them to become self-learners in biology, research study design, and the application of statistics to answer research questions in the biomedical sciences. The course is the first in a two-part, required sequence in the practice of biostatistics that students complete in their first year. The second course in the sequence covers the practical aspects of advising investigators on study design, writing analysis plans, conducting analyses, and reporting results.

2.2 Student Population

The course was designed for our two-year Master of Biostatistics program, which is housed within the School of Medicine and has the primary objective of training students to enter the workforce as highly effective collaborative biostatisticians. A secondary objective of the program is to prepare students who are interested in pursuing a PhD in Biostatistics or another related quantitative science. Historically, about 25% of students who complete our master's program will enter PhD programs after graduation. Admission to our master's program requires, among other things, a minimum of two semesters of calculus and one semester of linear algebra. Most of the students in this program are moving directly from undergraduate to graduate training, although some enter with a few years of work experience. We typically admit students who have undergraduate degrees in subjects like mathematics, biology, and public health. Students have the option of electing one of three specialty tracks in their second year: biomedical data science, mathematical statistics, or clinical and translational research. Our intent was to create a course that would be informative for students interested in any of these areas, and that would be positioned in the curriculum in the first year before students choose a specialty track.

3. Course Design

The course design process started in November 2017 and extended through July 2018, with the first iteration of the course offered in the Fall of 2018. The course design was accomplished by committee, led by the Director of Graduate Studies (DGS). Two active, collaborative researchers were selected as co-instructors for the course, one with expertise in biology and the other with expertise in biostatistics and research study design. The co-instructors were responsible for drafting a curriculum with input from the DGS and other, advisory committee members who represented expertise in specific areas of biostatistics that the curriculum was likely to cover, e.g., clinical trials. Committee meetings were held one to two times per month. Agreement on the course curriculum and instructional approach were arrived at through an informal, iterative consensus process.

The course consists of two core components: a 3-credit classroom experience and a required discussion section (see Appendix A for the course syllabus, and Appendix B for an example discussion section). The over-arching theme of the course is to explicitly link statistical methods to the biological basis for research. The approach is to introduce students to important concepts in research study design and biology using concrete examples from peer-reviewed literature using hands-on exercises that encourage critical thought, discussion, and application. The course design committee made a deliberate decision to avoid creating a course that provides comprehensive education in biology *per se*, but instead teaches students how to learn about the connection between biological sciences and statistics. Toward that end, the course is oriented around a single disease area, head and neck cancer, which serves as a

motivating example to: 1) introduce students to the most commonly used research study designs in medical and public health research, 2) illustrate how study design is used to address questions about human biology and disease, 3) teach basic biological concepts necessary for a successful career in biostatistics, and 4) train students to read and critically evaluate scientific research publications from peer-reviewed journals.

4. Course Features

4.1 Why Cancer?

Cancer was selected as a motivating example because of its public health importance, with cancer being the second leading cause of death in the United States (US) behind heart disease with ~600,000 deaths from cancer expected in the US during 2020 (Siegel et al., 2020). Cancer also represents a numerical enigma. For example, less than 1% of the US population is diagnosed with cancer each year, qualifying it as a "rare" disease. Yet, men and women of all ages in the US are faced with a 1 in 2 to 1 in 3 chance of being diagnosed with cancer during their lifetime, and there are currently over 15 million cancer survivors in the US (The National Cancer Institute, 2020). Cancer is also biologically puzzling. Cancer is fundamentally a disease of dysregulated cell growth, but it's occurrence in different tissues yields vastly different phenotypes and prognoses for patients, and the molecular mechanisms of cancer are even more vast and complex than its varied presentation (Allison & Sledge, 2014). There are environmental causes of cancer as well as endogenous causes (Ames et al., 1995). Cancer is typically a disease of old age, yet children are sometimes diagnosed (Bhakta et al., 2019). Some cancers are easily treated while others continue to present therapeutic challenges (Siegel et al., 2020). For all these reasons cancer is also a central organizing principle for understanding many central aspects of biomedicine, including molecular biology, genetics, cell biology, and evolution.

Head and neck cancers were selected because of the breadth of existing knowledge. Specifically, the disease has well-known etiology (Hashibe et al., 2009; Herrero et al., 2003), a well-researched molecular biology (Leemans et al., 2018), there are effective treatments (Lee et al., 2018), and excellent examples of all the major study designs exist in the literature. Because many students who take the course are unfamiliar with cancer, we ask students to read The Emperor of All Maladies (Mukherjee, 2010) as a gentle introduction to cancer from a layperson's perspective as part of their summer pre-orientation course, which they complete before taking this course(Neely et al., 2022). This book provides an excellent motivation for understanding the biological basis for research as it discusses many ground-breaking discoveries, as well as stunning failures, in the long history of cancer treatment. The first homework assignment in this course consists of open-ended questions about topics covered in the book that are intended to stimulate thinking about the link between biological sciences and statistical methods.

4.2 Teaching Students How to Think

The course materials are organized around a paradigm for evaluation of causal associations in medical research studies (Elwood, 2017). For example, this paradigm teaches students to evaluate medical evidence by first describing what they see (e.g., patient population, study design, stated hypothesis vs. tested hypothesis, and the main result) and then considering alternate explanations for the results (bias, confounding, etc.). In this way, students learn how to think deeply about medical evidence and avoid dogmatic approaches such as stereotyping clinical trial evidence "good" and evidence from observational studies as "bad". Under this paradigm of critical review, useful evidence can be generated by any study design and, likewise, any study design can be executed poorly.

The course provides students the opportunity to apply this "mental map" for evaluation of medical literature in various contexts, focusing on statistical thinking without doing much hands-on statistics (applied data analysis is covered in a second course the students are concurrently enrolled in). For example, concepts surrounding systematic error receive greater emphasis than random error, although the course includes a gentle conceptual introduction to the intersection of these two types of error in research studies. Finally, the approach is rooted in the use of published literature as examples. Importantly, both good and bad practices are highlighted in the examples.

4.3 Course Structure

The course is offered in the first semester of the two-year master's degree program. The classroom component requires active participation twice per week for 1.25 hours. Students are separated into small groups for the discussion sections, with each small group meeting every other week. The course follows the standard semester, which includes 13 weeks of instruction followed by a reading period and final exam. As shown in Figure 1, the course is composed of three integrated components: 1) research study designs commonly used in biomedical research, 2) biological concepts (essential biology, cancer biology, and specifics of head and neck cancers), and 3) critical

review of medical literature. The content in each of these components is delivered in blocks that are integrated such that biological concepts are taught in the context of study designs appropriate for addressing questions related to the biology.

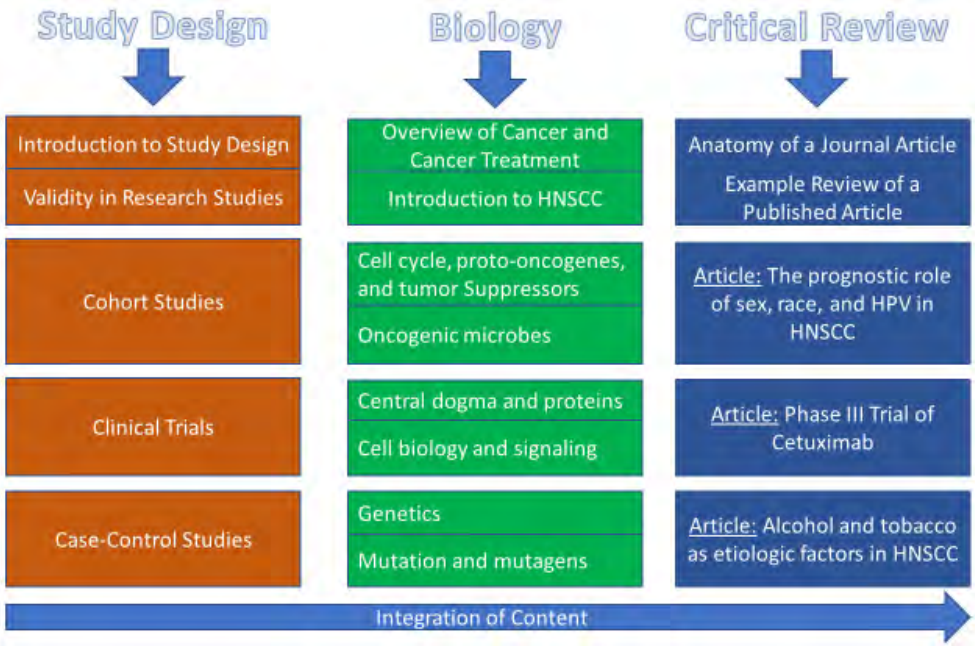


Figure 1. Design of a Course to Teach Biology and Communication to Graduate Students in Biostatistics Using Head and Neck Squamous Cell Carcinoma as an Example

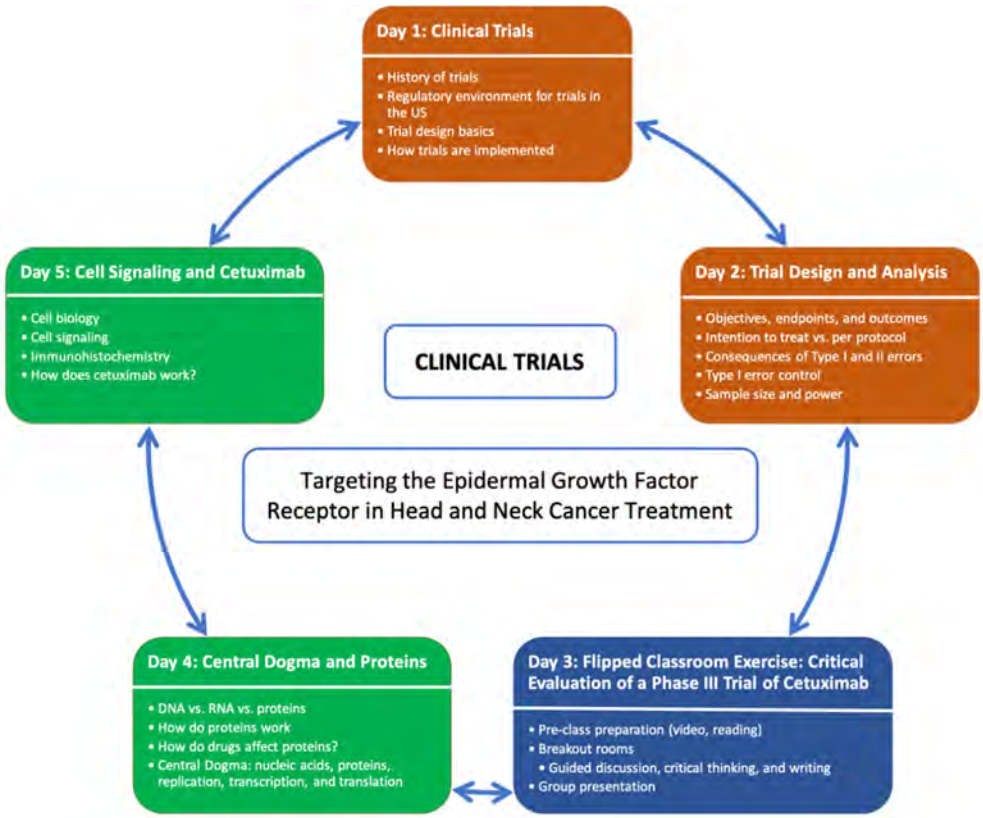


Figure 2. Example Course Module on Clinical Trials: Students Learn about Trial Design and How Drugs can Affect Proteins Using a Real-World Example of an Approved Drug for Treatment of Head and Neck Cancer

For example, as shown in Figure 2 and described in detail in Appendix C, the fundamentals of clinical trial design are taught simultaneously with concepts around proteins and cell signaling pathways as drug targets. The drug *cetuximab* is used as an example, which targets the epidermal growth factor receptor (EGFR) pathway in head and neck cancers. After covering clinical trial design concepts and the scientific basis for EGFR as a drug target, students read and critically review the primary publication from the Phase III trial leading to approval of cetuximab as a treatment for head and neck cancer. Using this approach, we can explicitly teach students how and why an in-depth understanding of the biological mechanism impacts study design and application of statistical methods.

A mixture of learning methods is used, including traditional didactic instruction, flipped classroom, and discussion sections. The mixture of learning methods serves to address some of the primary goals of the course surrounding development of communications skills. For example, the flipped classroom activity for the clinical trials module requires students to discuss a problem in small groups, write a solution to the problem, and present that solution to the class. This exercise emphasizes the importance of oral communication within the students' peer group (i.e., other statisticians), to an external audience, and provides an opportunity to practice concise scientific writing.

4.4 Evaluation of Student Performance

Student evaluation is based on homework assignments, participation in the discussion section, and written exams. Homework assignments and exam questions are open-ended to encourage critical thinking and to allow students the opportunity to practice scientific writing skills (see Appendix D for an example). Students can work together but are asked to write answers in their own words. The penultimate class activity is a group project that requires students to apply what they've learned to critically review research related to another disease and biological mechanism that was not covered during the class. This assignment allows students to apply what they have learned about study design and biology to a new scientific area, but also gives them the opportunity for self-discovery as they learn about a new topic on their own.

5. Student Feedback

Table 1. Proportion of Students Who Reported Being Satisfied or Very Satisfied on Core Survey Items by Year

	Survey Year (response rate)[a]		
	2018 (0.71)	2019 (0.88)	2020 (0.47)
Overall satisfaction with this course	13/20 (0.65)	19/23 (0.83)	21/22 (0.95)
Course Organization	9/20 (0.45)	13/23 (0.57)	18/22 (0.82)
Course Resources, Readings, and Materials	11/20 (0.55)	16/23 (0.70)	21/22 (0.95)
The discussions sections are integrated well with the classroom material	17/20 (0.85)	18/23 (0.78)	16/22 (0.73)
I learned a lot from the discussion sections	16/20 (0.80)	15/23 (0.65)	14/22 (0.64)

Numbers in the table are the number reporting being satisfied or very satisfied (numerator) over the number responding to the question (denominator) followed by the proportion in parentheses.

[a]The class size was 28, 26, and 47 students in 2018, 2019, and 2020 respectively. The class size was expanded in 2020 in accordance with the program's growth objectives. The 2020 class participated entirely online due to COVID-19.

In accordance with standard departmental practice, students are surveyed at the end of the semester prior to taking the final exam. Each year we made incremental improvements in the course in response to student feedback, e.g., by adding more homework assignments, incorporating a mid-term exam in addition to a final exam, and re-organizing the course website. Table 1 shows responses to the core survey questions from the first three offerings of the course. Students are asked to respond to a series of questions on a 5-point Likert scale with the top two rankings being Satisfied or Very Satisfied. In general, student satisfaction is high and has improved markedly over time. The one

area that we note for improvement relates to the discussion sections. While most survey respondents agreed that the discussion sections integrate well with the classroom component of the course, only ~2/3 of respondents reported satisfaction in 2019 and 2020 as compared to 80% reporting satisfaction with the discussion sections in 2018. The inaugural offering of the course in 2018 included 7 discussion topics whereas the course offerings in 2019 and 2020 included only 5 topics. One of the 2 discussions that were dropped was incorporated into the classroom component of the course, and the other (an exercise in which the students collaboratively designed a study during the discussion section) was dropped from the curriculum due to time constraints. This suggests that students may value some discussion topics more than others. In the future, we plan to ask for feedback with a series of targeted questions at the end of each discussion section in the future so that we might better calibrate the discussion topics to the needs of our students.

6. Summary

In this article we describe an innovative, graduate level course on biology and communications skills for biostatistics students. Our approach is highly contextualized, using critical review of medical literature in a well-researched area of human disease to teach core concepts and facilitate practice with oral and written communications. Importantly, our approach emphasizes how to think rather than what to know. We accomplish this by teaching students how to establish mental maps, or ways of thinking about problems using a set of common knowledge regarding study design and biology, combined with a paradigm for evaluating medical evidence. The course offers students the chance to learn through discussion and writing, which sharpens communication skills as they learn to think about complex scientific problems from the standpoint of a collaborative biostatistician.

Much has been written about the characteristics of successful collaborative biostatisticians that training programs should focus on (Begg & Vaughan, 2011; DeMets et al., 2006; Perkins et al., 2016; Van Steen et al., 2001). Nearly all emphasize the importance of excellent written and oral communication skills, and many have emphasized the need to be conversant in the clinical or biological domain within which the collaborative biostatistician is working (Pomann et al., 2020). Our course is a concrete example of how these core competencies can be taught. Although we offer the course to master's degree students, it could also be adapted to PhD programs with more depth to the content, e.g., by adding coverage of mathematical or theoretical foundations, or conducting computer simulations. Based on student feedback, our next steps in developing this course include upgrading the discussion sessions. Our next steps in overall curriculum development center around applying lessons learned in delivering this course throughout the curriculum.

In summary, the core focus of our course is to create lifelong learners who, when presented with unfamiliar scenarios will rely naturally on their critical thinking and communications skills to help them apply appropriate statistical methods for a given biological problem.

Acknowledgement

We would like to thank Terry Hyslop, PhD, Karen Pieper, PhD, and Beth Hauser, PhD for their advice during the design of this course.

References

Allison, K. H., & Sledge, G. W. (2014). Heterogeneity and cancer. In *ONCOLOGY (United States)*.

Ames, B. N., Gold, L. S., & Willett, W. C. (1995). The causes and prevention of cancer. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(12), 5258-5265. https://doi.org/10.1073/pnas.92.12.5258

Begg, M. D., & Vaughan, R. D. (2011). Are biostatistics students prepared to succeed in the era of interdisciplinary science? (and how will we know?). *American Statistician*, *65*(2), 71-79. https://doi.org/10.1198/tast.2011.10222

Bhakta, N., Force, L. M., Allemani, C., Atun, R., Bray, F., Coleman, M. P., Steliarova-Foucher, E., Frazier, A. L., Robison, L. L., Rodriguez-Galindo, C., & Fitzmaurice, C. (2019). Childhood cancer burden: a review of global estimates. *The Lancet Oncology, 20*(1), e42-e53. https://doi.org/10.1016/S1470-2045(18)30761-7

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, *32*(3), 347-364. https://doi.org/10.1007/BF00138871

DeMets, D. L., Stormo, G., Boehnke, M., Louis, T. A., Taylor, J., & Dixon, D. (2006). Training of the next generation of biostatisticians: A call to action in the U.S. *Statistics in Medicine*, *25*(20), 3415-3429 https://doi.org/10.1002/sim.2668

Elwood, M. (2017). *Critical Appraisal of Epidemiological Studies and Clinical Trials* (Vol. 1). Oxford University Press. https://doi.org/10.1093/med/9780199682898.001.0001

Hashibe, M., Brennan, P., Chuang, S. C., Boccia, S., Castellsague, X., Chen, C., Curado, M. P., Maso, L. D., Daudt, A. W., Fabianova, E., Fernandez, L., Wünsch-Filho, V., Franceschi, S., Hayes, R. B., Herrero, R., Kelsey, K., Koifman, S., Vecchia, C. La, Lazarus, P., … Boffetta, P. (2009). Interaction between tobacco and alcohol use and the risk of head and neck cancer: Pooled analysis in the international head and neck cancer Epidemiology consortium. *Cancer Epidemiology Biomarkers and Prevention*, *18*(2), 541-50. https://doi.org/10.1158/1055-9965.EPI-08-0347

Herrero, R., Castellsagué, X., Pawlita, M., Lissowska, J., Kee, F., Balaram, P., Rajkumar, T., Sridhar, H., Rose, B., Pintos, J., Fernández, L., Idris, A., Sanchez, M. J., Nieto, A., Talamini, R., Tavani, A., Bosch, F. X., Reidel, U., Snijders, P. J. F., … Franceschi, S. (2003). Human papillomavirus and oral cancer: The international agency for research on cancer multicenter study. *Journal of the National Cancer Institute*, *95*(23), 1772-83. https://doi.org/10.1093/jnci/djg107

Lee, Y. S., Johnson, D. E., & Grandis, J. R. (2018). An update: emerging drugs to treat squamous cell carcinomas of the head and neck. In *Expert Opinion on Emerging Drugs*, *23*(4), 283-299. https://doi.org/10.1080/14728214.2018.1543400

Leemans, C. R., Snijders, P. J. F., & Brakenhoff, R. H. (2018). The molecular landscape of head and neck cancer. *Nature Reviews Cancer*, *18*(10), 662. https://doi.org/10.1038/nrc.2018.11

Mukherjee, S. (2010). *The Emperor of All Maladies*. Scribner.

Neely M. L., Troy J. D., Gschwind G. T., Pomann G. M., Grambow S. C., & Samsa G. P. (2022). Preorientation Curriculum: An Approach for Preparing Students with Heterogenous Backgrounds for Training in a Master of Biostatistics Program. *Journal of Curriculum and Teaching, 11*(4), 77-85. https://doi.org/10.5430/jct.v11n4p77

Perkins, S. M., Bacchetti, P., Davey, C. S., Lindsell, C. J., Mazumdar, M., Oster, R. A., Peduzzi, P. N., Rocke, D. M., Rudser, K. D., Kim, M., Carter, R. E., Clinic, M., Cnaan, A., Geller, N. L., Goldberg, J. D., Mayo, M. S., Cayetano, S. M., Nietert, P. J., Pollock, B. H., … Welty, L. J. (2016). Best Practices for Biostatistical Consultation and Collaboration in Academic Health Centers. *American Statistician*, *70*(1), 7-30. https://doi.org/10.1080/00031305.2015.1077727

Pomann, G.-M., Boulware, L. E., Cayetano, S. M., Desai, M., Enders, F. T., Gallis, J. A., Gelfond, J., Grambow, S. C., Hanlon, A. L., Hendrix, A., Kulkarni, P., Lapidus, J., Lee, H.-J., Mahnken, J. D., McKeel, J. P., Moen, R., Oster, R. A., Peskoe, S., Samsa, G., … Thomas, S. M. (2020). Methods for training collaborative biostatisticians. *Journal of Clinical and Translational Science*, *5*(1), E26. https://doi.org/10.1017/cts.2020.518

Samsa, G. P. (2018). A Day in the Professional Life of a Collaborative Biostatistician Deconstructed: Implications for Curriculum Design. *Journal of Curriculum and Teaching*, *7*(1), 20-31. https://doi.org/10.5430/jct.v7n1p20

Samsa, G. P., LeBlanc, T. W., Locke, S. C., Troy, J. D., & Pomann, G.-M. (2018). A Model of Cross-Disciplinary Communication for Collaborative Statisticians: Implications for Curriculum Design. *Journal of Curriculum and Teaching, 7*(2), 1-11. https://doi.org/10.5430/jct.v7n2p1

Siegel, R. L., Miller, K. D., & Jemal, A. (2020). Cancer statistics, 2020. *CA: A Cancer Journal for Clinicians*, *70*(1), 7-30. https://doi.org/10.3322/caac.21590

The National Cancer Institute. (2020). *Cancer Statistics*. Retrieved from https://www.cancer.gov/about-cancer/understanding/statistics

Troy, J. D., Neely, M. L., Grambow, S. C., & Samsa, G. P. (2021). The Biomedical Research Pyramid: A Model for the Practice of Biostatistics. *Journal of Curriculum and Teaching*, *10*(1), 10-17 https://doi.org/10.5430/jct.v10n1p10

Van Steen, K., Curran, D., & Molenberghs, G. (2001). Towards a curriculum for the consultant biostatistician: Identification of central disciplines. *Statistics in Medicine, 20*(24), 3921-3929. https://doi.org/10.1002/sim.1082

Zapf, A., Huebner, M., Rauch, G., & Kieser, M. (2019). What makes a biostatistician? *Statistics in Medicine*, *38*(4), 695-701. https://doi.org/10.1002/sim.7998

Appendix

Appendix A: Course Syllabus

The following syllabus was used for the Fall 2020 course which was held during the COVID-19 pandemic. University and departmental policy required all learning to be remote and so the syllabus includes instructions for participation in live class sessions using Zoom.

Course Overview

This course will introduce biostatistics students to important concepts in research study design and biology using concrete examples of research studies published in peer-reviewed literature. The course uses *head and neck cancer* as a motivating example to: 1) Introduce students to the most commonly used research study designs in medical and public health research; 2) Illustrate how study design is used to address questions about human biology and disease; 3) Teach basic biological concepts necessary for a successful career in biostatistics; and 4) Train students to read and critically evaluate scientific research publications from peer-reviewed journals.

Class Format and Instructions for Participation

The course consists of a lecture and discussion lection. The lecture meets twice per week and each student attends a discussion section every other week. Students are assigned to a discussion section by the instructors. All lectures and discussion sections will be held virtually *via* Zoom at scheduled times. Attendance at all sessions is required unless extenuating circumstances prevent active participation. All sessions will be recorded. Use the course schedule (below) as a guide. Links are provided to the Sakai site where all course materials are kept.

Note that the class is divided into four discussion sections such each student attends a discussion every other week. You will either have a Tuesday afternoon section or a Friday morning section. Assignments to Tuesday or Friday are based on your time zone. Instructions for each discussion will be posted on Sakai in advance. You will have to spend some time preparing for each discussion, usually by reading an article and/or answering some questions. Everyone is expected to participate in the discussion, and participation is part of your course grade.

Most of the class sessions are didactic. A flipped classroom approach will be used for some sessions, particularly those that cover the example study designs – cohort, clinical trial and case-control studies. For these sessions you will have to do some work outside of class to prepare for an in-class discussion. In addition, during the last few weeks of class, we will switch from a didactic lecture format to an active learning exercise in which you work with your classmates to complete a case study and present the results to the class.

Homework Assignments

There will be 4 homework assignments. In general, the homework assignments will cover the topics presented in the 2-3 lectures preceding the date the homework is assigned. The only exception to this is the third homework, administered approximately in the middle of the semester, which will serve as a cumulative mid-semester review that is weighted more heavily than the other 3 homework assignments. You may work together on all homework assignments and use the textbook or online resources to answer the questions. However, you must write your answers in your own words. Do not copy or quote phrases from the textbook or any other resource.

Final Exam

The final exam for this course will be a take-home exam. You may work with your fellow students on this exam but the answers you turn in must be in your own words. Although you may use textbooks or online resources to answer the exam questions, you must write your answers in your own words. As with the homework assignments, do not copy or quote phrases from the textbook or any other resource.

Textbooks

Required: *Critical Appraisal of Epidemiological Studies and Clinical Trials*. An electronic version of this book is available for free from the Medical Library. To access the textbook, you must either be on campus or signed into the Virtual Private Network. The e-book URL is:

http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780198529552.001.0001/acprof-9780198529552

Recommended: *The Cartoon Guide to Biology* by Larry Gonick and David R Wessner (ISBN-10: 0062398652). Yes we are serious! This book is excellent background reading for all of the basic biology topics we will cover. Unfortunately, the *The Cartoon Guide to Head and Neck Cancer* is not out yet so we will depend on other sources for cancer specific readings. We will recommend specific sections of this book to accompany most of the biology lectures. If you understand biology at the level that is covered in this book you should well in the biology portions of

this course and the masters qualifying exam, and you will feel comfortable discussing the basic biology of biostatistical projects with collaborators who are biologists and physicians. If you took an introductory biology course as an undergraduate, you might not need this book.

Evaluation:

Your course grade will be determined from the following components.

- Homework assignments (30%)
- Discussion section participation (20%)
- Case studies (20%)
- Final exam (30%)

Grades are assigned as follows.

- A+ 96 – 100
- A 92 – 95
- A- 90 – 91
- B+ 86 – 89
- B 82 – 85
- B- 80 – 81
- C+ 76 – 79
- C 72 – 75
- C- 70 – 71

D - below 70

Course Schedule

All Classes and Discussion Sections are 1.25 Hours in Length. This Example Shows the Class Split into Four Groups—A, B, C, and D—for the Discussion Sessions (highlighted in orange) such that Discussions are Held in Small Groups. Green Highlighting Indicates When Homework is Assigned.

	Day	Topic
Week 1	1	Introduction to Head and Neck Cancer
		Study Design I
		Emperor of Maladies Questions Assigned
	2	Study Design II
Week 2	1	Overview of Cancer
		Discussion 1: Anatomy of a Paper – Group A
	2	Validity in Research Studies I
		Homework 1 assigned (covering lecture material from weeks 1-2)
	3	Discussion 1: Anatomy of a Paper – Group B
Week 3	1	Validity in Research Studies II
		Discussion 1: Anatomy of a Paper – Group C
	2	Evolution, Cancer Treatment, and Head and Neck Cancer
	3	Discussion 1: Anatomy of a Paper – Group D
Week 4	1	Cohort Studies I
		Discussion 2: The Biology of Serum Albumin as a Prognostic Factor in Head and Neck Cancer (Lim, et al.) - Group A
	2	Cohort Studies II
	3	Discussion 2: The Biology of Serum Albumin as a Prognostic Factor in Head and Neck Cancer (Lim, et al.) - Group B
Week 5	1	The Cell Cycle
		Discussion 2: The Biology of Serum Albumin as a Prognostic Factor in Head and Neck Cancer (Lim, et al.) - Group C

	Day	Topic
	2	Oncogenic Microbes
		Homework 2 assigned (covering lecture material from weeks 3, 4 and 5)
	3	Discussion 2: The Biology of Serum Albumin as a Prognostic Factor in Head and Neck Cancer (Lim, et al.) - Group D
Week 6	1	Example Cohort study: to be announced (flipped classroom)
		Discussion 3: The Biology of HPV in Head and Neck Cancer - Group A
	2	Clinical Trials I
	3	Discussion 3: The Biology of HPV in Head and Neck Cancer - Group B
Week 7	1	Clinical Trials II
		Discussion 3: The Biology of HPV in Head and Neck Cancer - Group C
	2	Example Clinical Trial - Phase III Trial of Cetuximab (flipped classroom)
	3	Discussion 3: The Biology of HPV in Head and Neck Cancer - Group D
Week 8	1	The Central Dogma and Proteins
		Discussion 4: Critical Appraisal of a Cohort Study (Fahkry, et al.) - Group A
	2	Cell Biology, Cell Signaling and Cetuximab
		Homework 3 assigned (mid-semester review covering weeks 1-8)
	3	Discussion 4: Critical Appraisal of a Cohort Study (Fahkry, et al.) - Group B
Week 9	1	Case-control Studies I
		Discussion 4: Critical Appraisal of a Cohort Study (Fahkry, et al.) - Group C
	2	Case-Control Studies II
	3	Discussion 4: Critical Appraisal of a Cohort Study (Fahkry, et al.) - Group D
Week 10	1	Example Case Control Study: A Case-control Study of Tobacco and Alcohol as Risk Factors for Head and Neck Cancer (flipped classroom)
		Discussion 5: Mechanism of Action of Cetuximab -Group A
	2	Genetics
	3	Discussion 5: Mechanism of Action of Cetuximab -Group B
Week 11	1	Mutation
		Discussion 5: Mechanism of Action of Cetuximab -Group C
	2	Case study 1: Total vs. Radical Mastectomy for Breast Cancer
		Homework 4 assigned (covering Weeks 9, 10 and 11)
	3	Discussion 5: Mechanism of Action of Cetuximab -Group D
Week 12	1	Case study 1: Total vs. Radical Mastectomy for Breast Cancer
	2	Case study 2: Pathological diagnosis of HPV-positive Head and Neck Cancer
Week 13	1	Case study 2: Pathological diagnosis of HPV-positive Head and Neck Cancer
	2	Review Session for the Final Exam

Appendix B: Example Discussion Section

The following is the first discussion topic of the semester. The purpose is to orient students toward reading medical journal articles.

Discussion 1: The Anatomy of a Medical Journal Article

Read the International Committee of Medical Journal Editors description of a journal article. Then, briefly review these other resources:

- The Laryngoscope's website
- Interpretation of the Impact Factor
- The Journal Citation Report for *The Laryngoscope*. *If for some reason the link I give you here doesn't work then go to https://jcr.clarivate.com and enter Laryngoscope into the Search box.*

Finally, read the title page and the Abstract for the article by Lim, et al. titled *Pretreatment Albumin Level Predicts Survival in Head and Neck Squamous Cell Carcinoma*. This article is posted on Sakai on the Discussion Sections page.

Be prepared to discuss answers to the following questions during the discussion.

1. The article is published in a journal called *The Laryngoscope.* Who is the primary audience (readership) for this journal? Where did you look to find this information?

2. How reputable is this journal? How do you know?

3. Now look at the study by Lim, et al. What does the title of the article imply about the study? Are any elements of the ICMJE recommendations missing from the title?

4. What are the authors' medical/research specialties and where are they from? Why do you think this information matters?

5. How was the research paid for? Why does this matter?

6. Evaluate how well the Abstract complies with the ICMJE recommendations for what should be present in an article Abstract. Are you able to find all of the elements? Is anything missing?

a. Context or background for the study

b. The study's purpose

c. Basic procedures (selection of subjects, participants, setting, measurements, and analytical methods)

d. The main findings of the study

e. The primary conclusions of the research, emphasizing the new information the study has revealed

f. Important limitations of the study

Appendix C: Example Module on Clinical Trials

The following outline relates to Weeks 6-8 in the example course schedule. The content covers an introduction to clinical trials in the context of drug development for treatment of head and neck cancer. The example uses cetuximab, an FDA-approved drug for treatment of head and neck cancer, as a tool for teaching students how clinical trials are designed to test biological hypotheses. There are two lectures each on clinical trial design and biology, as well as a flipped classroom session that requires students to read and discuss the published report of a Phase III trial of cetuximab.

After completing this module of the course students will be able to demonstrate the following knowledge related to clinical trials and cellular/molecular biology:

1. Discuss the regulatory environment for drug development in the United States.

2. Describe the typical phases of drug/device development.

3. Understand the commonly used parallel group clinical trial design.

4. Know the difference between objectives, endpoints, and outcome measures.

5. Define blinding (or masking) and understand why it is needed in clinical trials.

6. Understand the purposes of randomization in the context of treatment selection bias.

7. Describe how blocked and stratified randomization work, and why they are used.

8. Discuss how randomization might be implemented in a real study.

9. Know the difference between intention-to-treat and per-protocol analyses.

10. Discuss the motivations for addressing Type I error in trials and identify the pros and cons of some common approaches to Type I error control in trials with multiple endpoints.

11. Explain the roles of DNA, RNA, and proteins in cells, and how information flows among these macromolecules.

12. Give examples of some proteins and what they do.

13. Explain what a receptor is, what a ligand is, and their roles in cancer

14. Explain what Cetuximab is and how it works to treat cancer

The following readings are assigned for this module. Students may complete these at their own pace but are encouraged to read in parallel with the delivery of the course content.

1. Elwood, M. Critical Appraisal of Epidemiological Studies and Clinical Trials, 3rd Edition.

- P. 170-179 (Randomization)
- P. 130-132 (Single and Double Blind Methods)

2. Friedman, Furberg, and DeMets. Fundamentals of Clinical Trials, 5th Ed.
   - Chapter 1 (Introduction to Clinical Trials)
   - Chapter 3 (What is the Question?)
   - Chapter 5 (Basic Study Design), p. 89-95
   - Chapter 6 (The Randomization Process), p. 123-131
3. Gleevec Inhibits Cancer-Causing Kinase BCR-ABL
4. DNA Replication
   - Basic DNA Replication
   - Advanced DNA Replication
5. Transcription
   - Basic DNA Transcription
   - Advanced DNA Transcription
6. Translation
   - Basic Translation
   - Advanced Translation
7. Gonick, L. The Cartoon Guide to Biology.
   - Chapter 9: Meet the Genome
   - Chapter 4: Into the Cell

Clinical Trials, Day 1: Outline

Introduction to Clinical Trials

1. A brief history of clinical trials
2. Human subjects research in the United States: The Common Rule
3. Organization of the Department of Health and Human Services
4. The Food and Drug Administration: mission and organization
5. The phases of drug/device development
6. Clinical trial design basics: The parallel group trial
7. Randomization
   a. Goals of randomization
   b. Example of biased treatment selection
   c. What's random and what's not?
   d. Simple vs. restricted randomization
   e. Permuted block randomization
   f. Stratified randomization
8. How randomization is implemented in a real trial
   a. Blinding and allocation concealment
   b. Generation of randomization table using program code
   c. Assignment of treatments through an electronic system
9. Brief discussion of other treatment assignment strategies (e.g., minimization)

Clinical Trials, Day 2: Outline

Important Topics in the Design and Analysis of Clinical Trials

1. Specifying outcomes in clinical trials
    a. Objectives vs. endpoints vs. outcome measures
2. Analysis of clinical trials
    a. Intention to treat vs. per-protocol
    b. Brief introduction to missing data
    c. Hypothesis testing in clinical trials
        i. Consequences of Type I and II errors
        ii. Approaches to Type I error control
3. Sample size calculation for clinical trials

Clinical Trials, Day 3: Flipped Classroom

Description of the Activity

Before coming to class students read the article titled Radiotherapy plus Cetuximab for Squamous-Cell Carcinoma of the Head and Neck (Bonner JA, Harari PM, Giralt Jordi, et al. *N Engl J Med* 2006; 354:567-578) and watch a short video prepared by the instructors. The video is intended to guide the students through some of the more difficult parts of the reading assignment, e.g., by explaining in more detail what immunohistochemistry is and how it works. Slides are provided along with the video.

Students are given the following questions to review in advance of the class. Students are told that when they arrive for the class session they will be randomly divided into the breakout rooms. Students are given 20 minutes to discuss and write answers for their assigned questions. During the activity instructors travel between the breakout rooms to provide assistance and guide the discussion as necessary. At the end of the 20-minute period the class reconvenes, and students present the answers to the questions in front of the class. Some students nominate a single person from their breakout room to make the presentation, and in other cases multiple students present answers. The choice of how this is done is left entirely to the students.

Questions for the Breakout Rooms

Breakout Room 1. Apply Elwood's framework to describe the evidence presented in the article by Bonner (see page 324 of your textbook, Ex. 9.1, Part A).

a. Give a brief description of the intervention, including what it is, how it is expected to work, and how it is administered to patients.
b. What was the outcome? How was it measured or determined?
c. The study is described as a Phase 3 clinical trial. What is a Phase 3 trial?
d. Who was eligible for the study and where were participants recruited or identified from?
e. What was the main result of the study?

Breakout Room 2. Answer the following questions about blinding in this study.

a. Define what is meant by blinding in clinical trials and why it is important (this is a general question and not specifically about this trial)
b. How was blinding achieved in this study?
c. Give an alternate suggestion for how the investigators might have implemented blinding and discuss why you think the investigators did not take this approach.

Breakout Room 3. Table 2 shows a comparison of EGFR immunostaining by assigned treatment. Answer the following questions.

a. Describe the result shown in the table for EGFR

b. Why is it important to make this comparison? *Hint: think about the role the EGFR variable could potentially play in the study design given the investigator's biological hypothesis.*

c. Why is the Unknown category shown in the table? What must you assume about the Unknown group in order to accept the result shown in the table?

Breakout Room 4. The locoregional control endpoint was evaluated on an intention to treat basis.

a. Explain what "intention to treat" means

b. How many patients were included in this analysis, and of that number how many followed the protocol with respect to the use of radiation and cetuximab?

c. Why did the authors not focus the analysis only on those patients who complied perfectly with the protocol?

Breakout Room 5. Answer the following questions about the results for the locoregional control outcome.

a. Describe the result presented in Figure 1. Give both a description of what you see in the graph and an interpretation of the hazard ratio and 95% confidence interval discussed in the footnotes to the graph.

b. The authors give a long description of compliance with the assigned treatments in the Results section and in the Discussion the authors say the following: "The superiority of the cetuximab-plus-radiotherapy regimen we used cannot be attributed to underperformance in the radiotherapy group." Explain what this means and why it is important to mention.

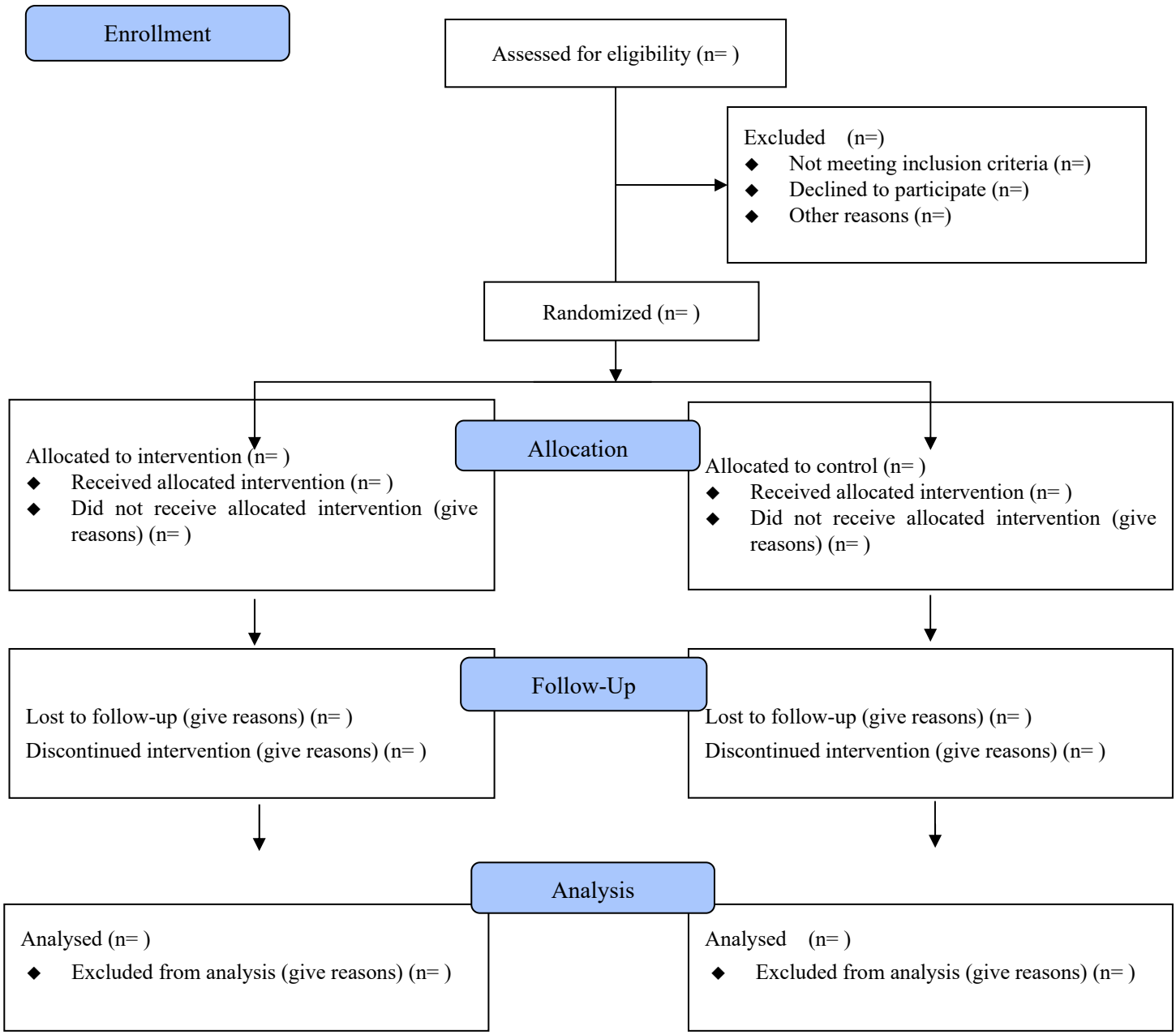
c. Patients were excluded from this study if they previously had cancer, received chemotherapy within the preceding three years, or had previously received surgery or radiotherapy for head and neck cancer. Why did the investigators decide to exclude such patients?

d. Briefly discuss internal and external validity of the results in terms of participant selection. Specifically, say whether and why you think the results apply to each of the following populations:

- Everyone who joined the study
- The eligible population
- The source population
- The target population

Breakout Room 6. Many clinical trials include a flow diagram that illustrates the design of the trial and the number of participants evaluated at each step in the study based on recommendations from the Consolidated Standards for Reporting Trials (CONSORT) organization. The article by Bonner did not include this type of diagram. Try to fill in the diagram based on the information given in the article.

CONSORT Flow Diagram



Clinical Trials, Day 4: Outline

The Central Dogma and Proteins

1. Introduction
    a. How does genetic information flow through a cell?
    b. What are DNA, RNA, and proteins?
    c. How are these molecules made? and what is their role in the cell?
    d. How do proteins work?
    e. How do drugs affect proteins?
2. Motivating example
    a. Gleevec for treatment of chronic myelogenous leukemia with the BCR-ABL fusion gene
3. Central dogma
4. Nucleic acids
5. Proteins
6. Replication
7. Transcription
8. Translation

Clinical Trials, Day 5: Outline

Cell Biology, Cell Signaling and Cetuximab

1. Cell biology
    a. Cell components
    b. Types of cells
2. Cell signaling
    a. Cell to cell signaling
    b. Paracrine, autocrine, and endocrine signaling
    c. Signal transduction
    d. Receptors and ligands
3. Overview of Cetuximab
    a. The epidermal growth factor receptor (EGFR)
    b. Antibodies
    c. Mechanism of action of cetuximab
4. Immunohistochemistry

Appendix D: Example Homework Assignment

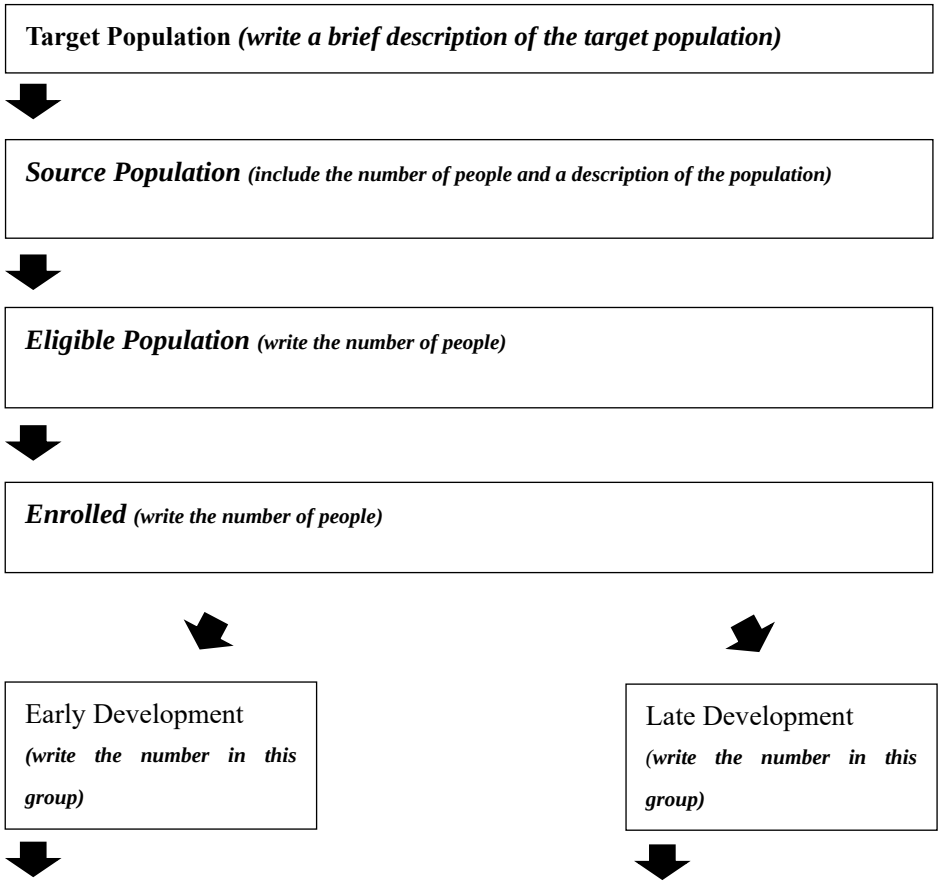
Homework 3: A Mid-semester Review Covering Weeks 1-8

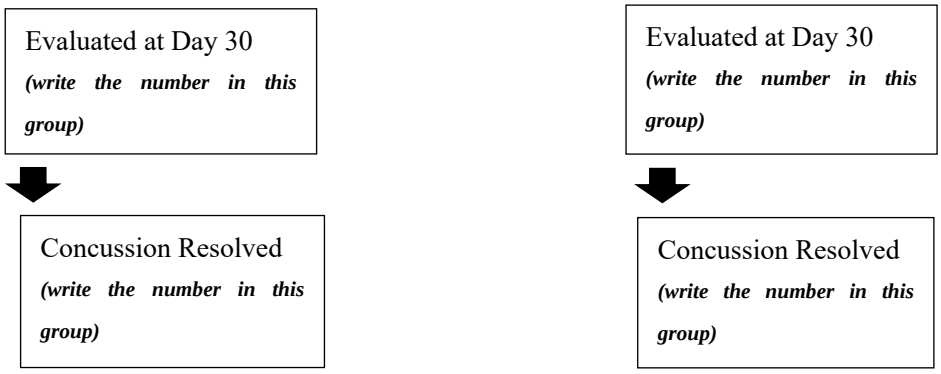
Instructions: Type short answers to the questions below (1-5 sentences). You may work on this assignment with your classmates, but your answers must be written in your own words.

Questions:

1. (**1 point**) Describe in your own words what the differences are between clinical trials at Phase I, II, III and IV in terms of research objectives, relative sample sizes, and types of endpoints evaluated at each phase.
2. (**1 point**) Randomization is effective for reducing the impact of biased treatment selection on study results. Explain in your own words what biased treatment selection is.

3. (**1 point**) Explain the purpose of stratified randomization and explain how it is different from stratified analysis.
4. (**1 point**) Describe the difference between a cohort study and a case-control study with respect to how eligible participants are selected.
5. (**5 points**) A prospective cohort study was conducted to evaluate the association between pubertal development and resolution of concussion symptoms in teenage ice hockey players. Recruitment for the study was conducted at 3 large, academic medical centers in the Northeast United States between September 1, 2012 and March 31, 2016. Patients visiting the outpatient sports concussion clinics at the 3 study centers during this time period (N=856) were eligible if they were age 13-19, had sustained a concussion while playing organized ice hockey in the week prior to the clinic visit, and were experiencing at least Mild symptoms. Patients were excluded if they had prior concussions in the last 6 months, or a history of 3 or more diagnosed concussions. A total of 302 eligible participants were offered participation in the study and 250 enrolled. Pubertal development was evaluated using the validated Pubertal Development Scale and patients were classified at baseline into two categories: Early (N=100) or Late (N=150) development. Patients were then followed for resolution of concussion symptoms, which was determined using the validated Post-Concussion Symptom Score, which measures severity of concussion symptoms on a 4-point ordinal scale (0=No symptoms, 1=Mild, 2=Moderate, 3=Severe). The primary outcome of the study was resolution of concussion symptoms at 30 days (Yes or No). A total of 10 participants in the early development group, and 15 participants in the late development group dropped out of the study prior to Day 30 and their outcomes were not measured. Among the 225 who completed the study, 90 were classified as Early Development and 135 as Late Development. At 30 days there were 53/90 (59%) with concussion resolved in the Early Development group and 35/135 (26%) in the Late Development group ($RR_{\text{early vs late}}$=2.27, 95% CI: 1.63 – 3.17).
    a. Write a brief summary of the primary result of the study, including an interpretation of the relative risk and 95% confidence interval.
    b. Fill in the following diagram depicting the selection of subjects for this study.





c. A stratified analysis was conducted to evaluate whether the observed association between pubertal development and resolution of concussion symptoms at 30 days was affected by the age of the participant. Participants were divided into two strata, Young (<15 years) and Old (>=15 years) for analysis. The table below shows the results of the analysis.

i. (1/2 point) What is the statistical phenomenon identified by the stratified analysis?

ii. (1/2 point) Write a short paragraph that summarizes the results of this stratified analysis, as if you were describing the results in a journal article. Include all relevant (based on your answer to (i) above) measures of association and confidence intervals in your description of the results.

	Young			Old		
	Concussion Resolved	Concussion Not Resolved	Total	Concussion Resolved	Concussion Not Resolved	Total
Early	3	29	32	50	8	58
Late	15	67	82	20	33	53
Total	18	96	114	70	41	111
RR (95% CI): 0.51 (0.16,1.65)				RR (95% CI): 2.28 (1.59,3.28)		
Breslow-Day Chi-square (1 degree of freedom)=6.20, critical value=3.84						
Mantel-Haenszel RR (95% CI): 1.78 (0.03,110.43)						

6. (**2 points**) The following refers to the ice hockey study described in question 5. Assume 10/25 participants who were lost to follow-up had resolution of concussion symptoms at Day 30 (4/10 of the losses from the Early group, and 6/15 losses from the Late group). What is this pattern of loss-to-follow-up called, and would you expect it to cause bias in the RR? If bias is present, in what direction would the bias be with respect to the value of the RR under the null hypothesis? Hint: try drawing a diagram similar to the ones we used during the lecture to visualize where participants drop out during follow-up.
7. (**1 point**) The process of making a protein based on RNA is called "translation." Why is it called translation? What is being translated?
8. (**1/2 point**) What is the primary function of DNA?
9. (**1 point**) Cetuximab is a standard treatment for colorectal cancer, however colorectal tumors with mutations in RAS are not treated with Cetuximab because it is generally not effective against these tumors. In normal cells, the EGFR protein activates the RAS protein by phosphorylation, and the active RAS protein then activates other proteins, ultimately promoting cell proliferation. How might mutations in RAS make Cetuximab ineffective?
10. (**1/2 point**) What is one type of RNA and its function?
11. (**1 point**) Most receptors have an extracellular part, a transmembrane part, and an intracellular part. The extracellular (outside of the cell) part is bound by the ligand. The intracellular part often effects the signal that the receptor has been bound by phosphorylating downstream proteins. The transmembrane part connects the extracellular and intracellular parts, allowing the binding signal to be transmitted from outside the cell to the inside, where it can alter cell activity. However, there are some receptors that lack

extracellular and transmembrane parts. How can these receptors receive signals from outside of the cell without an extracellular part? What is one signal that is detected by this type of receptor?

12. (**2 points**) The Central Dogma tells us that information from DNA is used to make three different types of polymers (molecules made up of multiple identical or similar subunits). Name two of these polymers and the key enzyme responsible for making each of these polymers.
13. (**1/2 point**) What is one protein and its function?