Abstract. Most undergraduates have limited experience with mathematical modeling. In an
effort to respond to various initiatives, such as the recommendations outlined in the Guidelines for
Assessment and Instruction in Mathematical Modeling Education (GAIMME) report [S. Garfunkel
and M. Montgomery, eds., GAIMME: Guidelines for Assessment & Instruction in Mathematical
Modeling Education, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2016], this
text describes a course on the mathematical models of cancer growth and treatment. Among the
aims is to provide a template for a “bridge” course between the traditional calculus and differential
equations sequence and more advanced courses in mathematics and statistics. Prerequisites for the
course include a course in ordinary differential equations. Linear algebra is a useful co-requisite but no
previous programming experience is required. The content includes classical models of tumor growth
as well as models for the growth of specific cancer types. Relevant research articles are provided
for further study. Material for student projects and effective communication is supplied, as well as
suggestions for homework assignments and computer labs. This paper aims to assist instructors in
developing their own “Mathematics + Cancer” course.

Key words. mathematical modeling, cancer, differential equations, undergraduate education

AMS subject classifications. 97M10, 92C50, 34A99

1. Introduction. This paper describes an undergraduate course, accessible to
students who have completed a standard sequence of calculus and ordinary differential
equations, on the mathematical modeling of cancer. The content and format of
the course are derived from the authors’ experiences in advising undergraduates in
program funded by the National Science Foundation’s Mentoring through Critical
Transition Points (MCTP) initiative. Our objectives in developing this course are
threelfold. First, we are interested in providing a model of a “bridge” course between
the traditional calculus sequence and higher-level courses besides the typical “intro-
duction to proof” class. Second, our effort is an attempt to develop an introductory
course in applied mathematics that addresses a compelling scientific and social prob-
lem. We motivate the relevant mathematical ideas at a level that is intelligible to a
broad student audience and in a way that will help students make informed choices
about more advanced courses in statistics, probability, numerical analysis, partial dif-
ferential equations, and dynamical systems, for example. Our third goal is to adapt
some of the pedagogical features of an undergraduate research experience—reading
papers from the primary research literature, completing a collaborative project, and
giving a talk—to a semester course format.

Our course also is an attempt to respond to recent programmatic initiatives
by professional mathematical societies, including those by Mathematical Associa-
tion of America’s (MAA) Committee on the Undergraduate Program in Mathematics
(CUPM), and by the Society for Industrial and Applied Mathematics (SIAM) and
the Consortium for Mathematics and Its Applications (COMAP). The 2015 CUPM

*This work was supported by the National Science Foundation’s Mentoring through Critical Transi-
tion Points program under grant number DMS-1148771.
†School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287-
1804. Current address: Department of Mathematics, University of Arizona, Tucson, AZ 85721
(stepien@math.arizona.edu).
‡School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287-
1804 (kostelich@asu.edu, kuang@asu.edu).
Curriculum Guide to Majors in the Mathematical Sciences [90] makes four “cognitive recommendations” for overall programmatic goals, stressing students’ development of communication skills, ability to apply theory to applications, facility with technological tools, and “mathematical independence and experience [of] open-ended inquiry.”

The Guidelines for Assessment and Instruction in Mathematical Modeling Education (GAIMME) report [39] by the SIAM and COMAP working groups discusses “transferable skills” that undergraduates can develop in the context of a modeling course, including identifying tractable questions, using reliable sources, working collaboratively, and communicating effectively. The Modeling Across the Curriculum report [24], which was funded by a National Science Foundation (NSF) grant to SIAM for “an initiative to increase mathematical modeling and computational mathematics in high school and college curricula” recommends developing accessible curriculum materials in addition to discussion of the modeling process.

Furthermore, by providing students with a research experience during a regular class, we are able to reach a diverse group of students that may not otherwise have the opportunity to participate in, for example, a supported project of the National Science Foundation’s Research Experiences for Undergraduates (REU) Program [77]. Many REUs are inaccessible to minority, first-generation, and/or nontraditional students, who for financial, logistical, or child-care reasons cannot attend an out-of-town program on a full-time basis for eight to twelve weeks. Our course represents an effort to provide a scalable, cost-effective alternative to a traditional REU. The enduring lessons the course aims to impress upon the students are similar to benefits students can obtain from participating in REUs: exposure to problem-solving experiences, awareness of STEM research fields and career options, and adding relevance to standard mathematics courses by applying theoretical knowledge to real-world cancer biology problems [72].

A final objective of this article is to motivate other efforts to develop courses with analogous goals on topics drawn from other areas of the mathematical sciences. We hope that the outline presented here, and the Supplementary Materials, may serve as a useful template.

2. Course Format and Overview. The course that we have developed meets for two 75-minute or three 50-minute periods per week over a 15-week semester. The first ten weeks or so use a “lecture-lab” format in which the main mathematical ideas are covered in a traditional lecture (one 75-minute or two 50-minute periods each week) followed by an informal computer lab whose exercises invite students to explore the ideas numerically. Homework sets, typically spaced every other week, emphasize the theoretical and analytical aspects in more detail. The last five weeks of the course are devoted to student projects and presentations. Given the limits of time and prior student experience, projects generally focus on reproducing the results of a mathematical model described in a research paper (Section 4).

Insofar as differential equations are used extensively to model the progression of cancer, students are assumed to have had a prior course in ordinary differential equations, such as might follow the second or third semester of calculus. (Linear algebra is a useful co-requisite.) No previous programming experience is required. We have used MATLAB [100] for the computer lab exercises, but GNU Octave [41], Maple [67], Mathematica [110], and Python [22] are potential alternatives. Any platform that allows for interactive use and that includes an easy-to-use two-dimensional plotting facility should suffice.

Because one of our pedagogical goals is to give students experience in reading
research papers, we have not used a traditional textbook. Nevertheless, instructors may wish to consider portions of texts by Wodarz and Komarova [109] and Kuang, Nagy, and Eikenberry [61] as background material for lectures. Our course so far has used ovarian cancer, prostate cancer, and glioblastoma multiforme as representative case studies, in part because of our research connections with oncologists specializing in these cancers. Table 1 lists some papers on these topics and others that, in the authors’ opinion, have been suitable for case studies and student projects.

Few of the students have previously read an article in a research journal, so we provide some coaching. Our approach is to liken a research paper to the layers of an onion: first read the abstract, introduction, and conclusions for the basic gist; skim the methods and analysis section; then try to understand more details on subsequent readings. Greenhalgh’s text, How to Read a Paper [44], is a useful supplement; the book focuses on the medical literature and is written for a general scientific audience.

Another goal of the course is to provide students with experience in communicating science beyond the written homework and blackboard problem presentations that are typical of first- and second-year college mathematics courses. We ask the students to articulate to their peers why their problem is important and interesting and to summarize the main findings in a short talk and/or poster format. Such training is common in graduate programs, but we think that nearly all undergraduates can benefit, insofar as analogous types of oral communication are necessary in most professional vocations.

3. Course Modules. In this section, we describe various modules of one to two weeks in length that can serve as subject material for lectures and computer labs. A sample syllabus, homework problem sets, and computer lab descriptions may be found in the Supplementary Materials.

Rarely a day goes by without newly published news articles on cancer, which can provide relevant material for class discussions. Cancer is the second most common cause of mortality in the United States (accounting for about one in every four deaths [5]). While cancer survival rates overall have increased in recent years, there has been no improvement for teenagers and young adults [21], for reasons that are unclear. New treatments are being tested constantly, and it is a challenge to stay current on the subject. We have tried to include one or two guest lectures by local research oncologists each term, as they can provide a clinical perspective of the challenges and legal and ethical obligations to patients when selecting standard and/or experimental treatments.

Since our course is offered within a mathematical sciences program, we focus on mathematical and statistical models of cancer. Nevertheless, we begin the course with an overview of cancer biology. Hanahan and Weinberg [46] describe six “hallmarks” of cancer growth: sustaining proliferative signaling, evading growth suppressors, activating invasion and metastasis, enabling replicative immortality, inducing angiogenesis (growth of blood vessels), and resisting cell death.

3.1. What and Why do We Model?. Our main purpose is to discuss the role that mathematics and data play in understanding the dynamics of cancer, from the standpoint of an individual tumor to the burden of cancer at the population level. What and why do we model? can focus on the following questions:

Statistical models: What fraction of people will be diagnosed with cancer in the next year? How can one account for the increasing incidence of cancer with age?

Dynamical models: How does a population of cancer cells grow, spread, and evolve
resistance to treatment in an individual patient?

**The scale problem:** What are the tumor dynamics at the cellular and tissue level?

**The data problem:** What can we measure? How do clinical observations, such as medical imaging, relate to the underlying disease process? Insofar as many tumors are incurable, how can one measure the efficacy of treatment?

**Policy:** From a public health perspective, how much cancer might be preventable (e.g., by anti-smoking and -obesity campaigns), how much is due to environmental exposures (e.g., sunlight, pollution), and how much is simply “bad luck” (i.e., arising from random somatic mutations)?

**Prediction:** Does a model using particular biological assumptions produce tumors that behave in a manner consistent with clinical observations? If not, then why not? Can we make short-term forecasts of tumor progression in an individual patient?

**Hypothesis testing:** How are statistics used to design clinical trials and to assess the efficacy of new treatments?

**In silico clinical trials:** How might the sequencing of treatment (e.g., surgery before radiotherapy) affect the growth of a tumor under various hypotheses?

Throughout the course, we return to these ideas repeatedly to evaluate the mathematical models that are discussed. To validate the model equations, dimensional analysis and analysis of qualitative behavior and end behavior is studied as well as errors and data fitting. These analyses can be done during lecture or in the homework and computer labs (suggested assignments are given in the Supplementary Materials). References such as [13, 29], among others, that describe principles of mathematical modeling can be incorporated into class discussions.

### 3.2. Statistical Modeling of Cancer

Statistical and probabilistic analyses of data have been used for many years to understand the genesis of tumors. Our classroom discussion begins with two classic papers from the 1950s to demonstrate the value of mathematical analyses to unravel the mystery of cancer formation.

Using the first reliable, comprehensive death registries in Western Europe and the U.S., Nordling [79] observed that the risk of dying from cancer was roughly proportional to the sixth power of age. If, for example, one crucial mutation sufficed to induce a cancerous tumor, and if there were an approximately constant probability $p$ that a given dividing cell would incur such a mutation, then the death rate from cancer would be roughly constant with age. Consequently, Nordling explained the observed power law relationship by suggesting that seven mutations were necessary, on average, for tumors to arise, assuming that the probability of the next mutation in the required sequence is approximately constant throughout life.

In 1954, Armitage and Doll [7] used death registries from England and Wales to examine the risk of cancer death by site as a function of age. They found that the death risk from colorectal, stomach, pancreatic, and esophageal cancers followed Nordling’s power law relationship reasonably well but that the death risk from lung, prostate, breast, ovarian, cervical, and uterine cancers did not. Armitage and Doll proposed a time-varying modification of Nordling’s hypothesis. If the $s$th mutation ($1 \leq s \leq 6$) required for a given cancer occurs by age $t_0$, and cancer is diagnosed at some later age $t$, then they argued that cancer risk should be proportional to $t_0^{-1}(t - t_0)^{6-s}$. Such a theory could explain the observed age-related incidence of cancers of the reproductive tract if the rate of initial mutations were a function of hormones or similar factor.

Armitage and Doll later refined their approach and proposed a two-stage theory.
A key public policy question concerns how cancer prevention and screening efforts should be prioritized. We focus one lecture on a review of Tomasetti and Vogelstein [104], which postulates that about 65 percent of the difference in lifetime cancer risk among various tissues can be explained by the average number of stem cell divisions within them. A subsequent paper [102], using data on 17 cancer types from 69 countries, corroborates the original findings and estimates the relative contributions of environmental, inherited, and random factors to the incidence of some common cancers. The authors’ mathematical analysis concludes that “primary prevention is the best way to reduce cancer deaths” but also that many cancers are not preventable [102], so screening efforts will remain necessary.

3.3. Experimental Design and Hypothesis Testing. A discussion of the mathematics behind disease testing and clinical trials introduces students to the complexities of determining the extent to which treatments may be regarded as successful. Much of the content in this part of the course is classical, but many students have not had previous exposure to it. One objective is to interest students in more advanced courses in probability and statistics.

We begin this portion of the course with basic set theory including union, intersection, and complement of sets as well as the Inclusion–Exclusion Principle. This introduction leads into a discussion of what a sample space, event, and probability are. The next concept is conditional probability, which is the probability of event $A$ occurring given that event $B$ has occurred, or

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

An important theorem that follows is Bayes’ theorem, which says that if $P(B) > 0$ and $P(B^c) > 0$ (where $B^c$ is the complement of $B$), then for any event $A$ such that $P(A) > 0$, then

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}.$$

One application of Bayes’ theorem occurs in disease screening. For example, suppose that a test correctly identifies 95% of the people who are known to carry a particular genetic mutation (i.e., the test’s sensitivity is $P(+|M) = 0.95$) and that the test also rules out 99% of the population that does not have the mutation (the specificity is $P(−|M^c) = 0.99$). If the mutation is rare (e.g., the prevalence is $P(M) = 0.001$), then (2) can be used to show that the probability that one carries the mutation given a positive test (the precision, $P(M|+) = \frac{P(+|M)P(M)}{P(+|M)P(M) + P(+|M^c)P(M^c)}$)

$$= \frac{0.95 \times 0.001}{0.95 \times 0.001 + 0.99 \times 0.999} \approx 0.087,$$

or about 8.7%, using the fact that $P(+|M^c) = 1 - P(−|M^c)$ and $P(M^c) = 1 - P(M)$. On the other hand, the probability that one does not carry the mutation given a negative test (the negative predictive value, $P(M^c|−)$) is greater than 99.9%. Examples
like this illustrate why, based on a negative test, a doctor may be able to reassure a
patient that he or she almost certainly does not have a given disease, and why another
test is necessary to confirm a diagnosis otherwise.

We also devote a couple of class periods to null hypothesis significance testing,
which is a standard topic in introductory statistics courses. Nevertheless, an intro-
ductive discussion of experimental design seems warranted, given the importance of
clinical trials to cancer treatment and research and because many of our students
have not yet had a course in statistics. We review the classical tests for differences
between sample means (e.g., between an experimental and a control group) and the
computation of associated confidence intervals. Students also read about some of the
criticisms of the classical approach (e.g., Cohen [20]) and common misconceptions
about p values (e.g., Goodman [43]), which lend themselves to interesting classroom
discussions. For example, in null hypothesis significance testing, the choice of the sig-
nificance level is arbitrary, the results depend on the sample size, and one can fail to
detect a treatment effect if the sample size is too small. Such underpowered trials are
a real-world problem and potentially an ethical one, insofar as patients may continue
to receive an inferior treatment. Furthermore, statistical significance does not imply
clinical significance, because the classical approach does not quantify the effect size.

There are many other statistical topics that can be developed in greater detail,
depending on the instructor’s interest. One of our computer labs introduces the
concept of bootstrap sampling [30], using an example from Efron and Tibshirani [31]
to compute confidence intervals for hazard ratios of heart attack and stroke risk among
men taking baby aspirin. Other topics might include selection and publication bias in
clinical trials, power analysis of trial designs, sensitivity analysis and quantification
of errors, ensemble forecasting, and the like.

Cancer data to use for statistical studies can be found on the Internet, for example
on web sites maintained by the Surveillance, Epidemiology, and End Results (SEER)
Program of the National Cancer Institute (NCI) [95], American Cancer Society [3],
and Centers for Disease Control and Prevention (CDC) [18].

3.4. Dynamical Models of General Tumor Growth. The ordinary differ-
ential equations emphasis in this course naturally starts with a discussion of growth
models that can be applied to tumor cell proliferation. To build students’ founda-
tions in mathematical oncology modeling, we introduce historically-relevant models
that have been used to analyze tumors’ uncontrolled cell proliferation. While students
learn about how the model equations are derived, it is useful to review important ba-
sic concepts like separation of variables, direction fields, equilibrium points and their
stability, the effect of initial conditions and parameters on the solution, and topics
related to systems of ordinary differential equations such as phase portraits and eigen-
values and eigenvectors which are best illustrated with simple models. We suggest
that this is done for each of the models we list below so that students see what is
qualitatively and quantitatively different among the equations and what that could
mean for tumor growth in the long run.

The clonal expansion of a mutant cell line, wherein one cell divides into two
daughter cells, leads to the difference equation $x_{n+1} = \alpha x_n$ with initial condition $x_0$
and growth rate parameter $\alpha$ and to the continuous exponential growth equation,

$$\frac{dx}{dt} = \alpha x,$$

with initial condition $x(t_0) = x_0$. The dependence of the doubling time on $\alpha$ can be
discovered and analyzed.
Of course, resource and space limitations imply that no cell population can grow without bound. The logistic equation
\[
\frac{dx}{dt} = \alpha x \left(1 - \frac{x}{K}\right)
\]
introduces the notion of a carrying capacity \(K\), as does the regrowth model,
\[
\frac{dx}{dt} = \alpha \left(1 - \frac{x}{K}\right).
\]
Both of these models are specific cases of cell growth rates given by the more general principle
\[
\frac{dx}{dt} = \text{growth rate} - \text{death rate}.
\]
The von Bertalanffy growth model [107] is derived from the observation that the pulse rate and metabolic rates \((M)\) in mammals are often proportional to the animal’s surface area \((SA)\) rather than to the volume \(V\) of the body. Since \(SA \propto V^{2/3}\), \(M \propto V^{2/3}\). More generally, one may assume that \(M \propto V^\lambda\) for some \(\lambda \in \left[\frac{2}{3}, 1\right]\) which leads to the growth model
\[
\frac{dx}{dt} = \alpha x^\lambda - \beta x^\mu,
\]
where \(\alpha\) and \(\beta\) are proportionality constants and the other parameters \(\lambda\) and \(\mu\) are positive. The case where \(\lambda = \mu = 1\) reduces to the exponential equation (4) and the case where \(\lambda = 1, \mu = 2\) reduces to the logistic equation (5). The case where \(\lambda = \frac{2}{3}, \mu = 1\) is called the surface area model [101].

The Gompertz growth model [42] arose from an actuarial accounting of the number of people alive as a function of their age. This model assumes that the population \(x(t)\) grows according to an exponentially decaying birth rate \(g(t)\),
\[
\begin{align*}
\frac{dx}{dt} &= g(t)x(t), \\
\frac{dg}{dt} &= -\alpha g(t).
\end{align*}
\]
Let \(g(0) = g_0\) be the initial population’s net fecundity. Substituting the solution of (9b) into (9a) yields the relation
\[
\ln x(t) = \ln x_0 + \frac{g_0 - g(t)}{\alpha},
\]
where \(x(0) = x_0\) is the initial population size. Since \(g(t) \to 0\) as \(t \to \infty\), it follows that \(\ln x(t) \to \ln x_0 + g_0/\alpha\) as \(t \to \infty\). If \(\lim_{t\to\infty} x(t)\) is interpreted as the carrying capacity \(K\), then \(x(t) \to K\) as \(t \to \infty\). Thus, \(g_0 = \alpha \ln (K/x_0)\), which implies that \(g(t) = \alpha \ln (K/x(t))\), so (9a) becomes the Gompertz growth model,
\[
\frac{dx}{dt} = \alpha x \ln \left(\frac{K}{x}\right).
\]
Other derivations of (10) are possible [61, 101]. An important early attempt to validate the Gompertz equation qualitatively with laboratory data can be discussed as

This manuscript is for review purposes only.
an introduction to calibration, inference, and prediction in the context of deriving, testing, and applying mathematical models of tumor cell proliferation [62, 63].

A discussion involving which model equations appear more physically-relevant on qualitative and quantitative levels can be held after analyzing the ordinary differential equations. Numerical simulations of these different growth functions can be explored in a computer lab such as described in the Supplementary Materials for quantitative comparison of solutions. Further investigation of these models is most appropriate when considering specific cancers (instead of general tumor cell proliferation models like in this section), as we do for the remainder of the course.

3.5. Mathematical Models of Chemotherapy and Surgery. After general tumor growth models are introduced, we then move into a discussion of how the effects of treatment can be described by a deterministic model. Surgery, chemotherapy, radiotherapy, and immunotherapy are examples of common treatments that can be included in mathematical models. One specific example is the ovarian cancer model of Kohandel et al. [58].

The American Cancer Society [5] estimates that there will be 22,530 new cases of ovarian cancer and 13,980 deaths in the United States in 2019. Most of these cases occur in women past menopause, and the most common type is ovarian epithelial cancer, accounting for approximately 95% of the cases. Currently, chemotherapy and surgery (laparotomy) comprise the two main treatments for Stage IIIC/IV ovarian cancer, in which the tumor is larger than 2 cm in diameter and has metastasized beyond the abdomen. A key question is: Does the order of treatments matter?

A recently concluded, large multicenter randomized clinical trial sought to study the differences in survival between surgery prior to chemotherapy versus chemotherapy prior to surgery [106]. While in clinical settings, the usual therapeutic strategy is to perform surgery before chemotherapy, it was found that the opposite sequence is not inferior. What can mathematical models tell us about the difference in the order of treatments?

Consider the generic model of tumor growth (7), where $x(t)$ is the tumor cell population at time $t$. The growth rate prior to treatment can be modeled by any of the right-hand sides of (4), (5), (8), or (10), denoted by function $f(x)$.

The death rate, denoted $P(t, x)$, describes the pharmacokinetic and pharmacodynamic effects of chemotherapy on the cancer. Various cell-kill hypotheses have been studied and are useful for students to compare and test for model validity, such as the log-kill hypothesis [91], in which chemotherapy kills cancer cells at a rate proportional to their population,

$$P(t, x) = cx,$$

the Norton-Simon hypothesis [80, 81], in which chemotherapy kills cancer cells at a rate proportional to their growth rate,

$$P(t, x) = cf(x),$$

and the $E_{\text{max}}$ model [49] in which chemotherapy kills cancer cells at a saturable rate,

$$P(t, x) = \frac{cx}{x + \delta}.$$ 

Treatment produces a response when the death rate is strictly larger than the growth rate.
To include the effects of surgery in the model, we observe that surgery affects the number of tumor cells, and thus the initial condition of the differential equation. Kohandel et al. [58] explore mathematical models of the two different treatment orders, termed adjuvant chemotherapy and neoadjuvant chemotherapy.

Adjuvant chemotherapy refers to the case where chemotherapy is given following the primary therapy (here, surgery, performed instantaneously at $t = t_s$). Suppose the tumor initiates at $t = 0$ and grows untreated until surgery. Then

\[
\begin{align*}
\frac{dx}{dt} &= f(x), & t &\in [0, t_s], \\
x(0) &= X_0,
\end{align*}
\]

where $f$ is the growth model. Assume that surgery removes all but $e^{-k}$ of the tumor at $t = t_s$; then the tumor population $x_c(t)$ after surgery under chemotherapy until final time $t_f$ is modeled by

\[
\begin{align*}
\frac{dx_c}{dt} &= f(x_c) - P(t, x_c), & t &\in [t_s, t_f], \\
x_c(t_s) &= e^{-k} \left( \lim_{t \to t_s^-} x(t) \right).
\end{align*}
\]

The final tumor size following treatment is

\[
x_{SC} = x_c(t_f).
\]

Neoadjuvant chemotherapy refers to the case where chemotherapy is the initial treatment, beginning at $t = t_c$ (surgery follows at $t = t_s$). Again supposing that the tumor initiates at $t = 0$ and grows untreated until chemotherapy begins, we have

\[
\begin{align*}
\frac{dx}{dt} &= f(x), & t &\in [0, t_c], \\
x(0) &= X_0,
\end{align*}
\]

and the treatment model becomes

\[
\begin{align*}
\frac{dx_c}{dt} &= f(x_c) - P(t, x_c), & t &\in [t_c, t_s], \\
x_c(t_c) &= x(t_c).
\end{align*}
\]

Surgery occurs at $t = t_s$ and again removes a fraction $e^{-k}$ of the remaining tumor, which yields a final tumor population of size

\[
x_{CS} = e^{-k} \left( \lim_{t \to t_s^-} x_c(t) \right).
\]

Which sequence of treatments yields the smallest tumor (or whether there is any difference) depends on the untreated tumor growth model $f(x)$ and the chemotherapy model $P(t, x_c)$. Analytical results can be obtained in some cases, and numerical methods must be employed in others depending on which cell-kill hypothesis is studied. This example lets students test the results of various assumptions about the growth and treatment models (which includes the option of having students create their own possible models) to determine which ones might be physically-relevant and to perform some parameter sensitivity experiments as outlined in the homework and computer labs in the Supplementary Material.
3.6. Mathematical Models of Prostate Cancer Treatment. Prostate cancer is one of the most common tumors in men. (The American Cancer Society estimates a 1 in 9 lifetime risk, with 174,650 new cases of prostate cancer in 2019 and 31,620 deaths [5].) During the 1930s and 1940s, surgeon Charles Huggins showed that castration (orchietomy) in men with locally metastatic prostate cancer usually causes dramatic regression of the disease. His results proved that most prostate tumors are hormone dependent (as are normal prostate cells [37]) and provided some of the first evidence that certain cancers were potentially controllable by chemical means. Similar results eventually were shown for most breast cancers [50]. Huggins shared the Nobel Prize in Medicine in 1966 for this discovery [4].

Castration remains the standard of care for locally advanced prostate cancer, although nowadays the treatment is usually pharmacological rather than surgical [94]. So-called androgen deprivation therapy (ADT) typically is highly effective, but it can cause serious side effects, including bone loss, diabetes, and heart and kidney problems [94]. Moreover, the tumor almost always evolves resistance to ADT, and there is no consensus regarding the choice and efficacy of subsequent treatment options [23]. ADT typically is given continuously until resistance occurs. Some clinical trials are investigating the use of intermittent therapy, whereby patients receive ADT until serum levels of a tumor marker (typically prostate-specific antigen, or PSA) drop below a predetermined level, at which point ADT is discontinued until the patient’s PSA levels rebound. The cycle continues until the tumor demonstrates resistance (e.g., PSA levels fail to drop sufficiently after a certain time interval). The hope is that intermittent ADT will reduce side effects and delay resistance [51, 1].

There are several challenges in devising mathematical models of ADT. First, one wishes to capture the initial rapid regression of the cancer in response to ADT, and second, to capture the evolution of resistance. Ideally, one would like to find a model that can quantify a typical patient’s treatment response, and, in the case of intermittent ADT, predict whether another cycle is likely to be effective for a given patient.

One model of the initial response of prostate cancers to ADT borrows from an ecological model of population growth under resource constraints, popularized by Droop [28]. In this view, there is a minimum cell quota, $q_{\text{min}}$, of some essential nutrient $Q$ below which the population $x(t)$ begins to shrink:

$$\frac{dx}{dt} = \mu \left(1 - \frac{q_{\text{min}}}{Q}\right) x.$$  

Here $q_{\text{min}}$ may be regarded as a physiological constant, and $\mu$ is a (positive) net proliferation rate. If $Q > q_{\text{min}}$, then the term in parentheses is positive and $x$ increases, but if $Q < q_{\text{min}}$ then $x$ decreases, possibly rapidly. Here we regard $x$ as the population of prostate cancer cells and $Q$ as the cells’ internal supply of androgen (or associated metabolite).

A very simplified model of ADT can be constructed by supposing that $Q$ is a time-varying function that depends on the present serum level of androgen, $A(t)$, which can be measured clinically (and is manipulated pharmacologically during treatment). We suppose that the tumor cells’ internal androgen supply grows according to the present serum androgen level and is consumed by the cellular metabolism at a net rate given by

$$\frac{dQ}{dt} = c_1 \left(\frac{A}{A + \nu}\right) - c_2 Q.$$  

This manuscript is for review purposes only.
The constants $c_1$, $c_2$, and $\nu$ are chosen so that, at pre-treatment levels of serum androgen $A(t)$, the right-hand side of Eq. (21) is approximately zero; initially, $Q(t_0)$ and $\mu$ are chosen to yield net growth of tumor that is positive but clinically realistic. When ADT reduces $A$ to a fraction of its pre-treatment level, $dQ/dt$ becomes negative, $Q(t)$ eventually drops below $q_{\text{min}}$, and the tumor responds accordingly. Treatment resistance might be modeled by supposing that $\nu$ is not constant but gradually declines, perhaps at the rate $\nu' = -\beta \nu$ for some positive constant $\beta$, which eventually renders $Q(t)$ independent of $A$.

A more sophisticated model by Portz, Nagy, and Kuang [83] postulates the existence of two tumor phenotypes, treatment-sensitive and -resistant, that depend on internal cell quotas of androgen in a Droop-like manner and that also can mutate from one phenotype to the other depending on treatment pressure. The model, which consists of five ordinary differential equations, also estimates the tumor’s production of PSA. It is possible to estimate parameters for this model that yield time series of PSA that are roughly consistent with those observed from individual patient cases in a clinical trial [1]. One of the most challenging computer labs in the course asks students to program the Portz-Nagy-Kuang model (with suitable constants) in MATLAB and to drive it using clinically observed levels of serum androgen, $A(t)$, in individual patients to reproduce some of the results in [83].

3.7. Mathematical Models of Glioblastoma Multiforme. The final case study in our course involves glioblastoma multiforme (GBM), which is the most common primary brain tumor in adults, and, unfortunately, has a poor prognosis [78]. The location of the tumor can affect patients’ neurological symptoms. Insofar as these tumors are highly diffusive as well as proliferative, a partial differential equation model is necessary. We include a brief introduction to the diffusion equation as the last mathematical topic in the course.

Typical models of GBM growth express the rate of population growth as the sum of a diffusive term and a proliferation term, which leads to a reaction-diffusion model of the form

\[
\frac{\partial u}{\partial t} = \nabla \cdot (D \nabla u) + g(u).
\]

In the simplest formulation, $D$ is a constant, and $g(u)$ is an exponential or logistic growth term [96, 97]. Normally, (22) is simulated in three dimensions on an anatomically accurate brain domain, but such a numerical implementation is too difficult for a course of this nature. Instead, we study (22) along an interval, which might be regarded as a model of diffusive migration of glioblastoma cells along a bundle of nerve fibers.

The usual diffusion equation (with $g(u) = 0$) permits a discussion of the classical methods of separation of variables for various initial and boundary conditions. GBM rarely metastasizes beyond the central nervous system [45] or penetrates the skull, so no-flux boundary conditions are appropriate for (22). Students simulate (22) numerically in the one-dimensional formulation using finite differences as a system of ordinary differential equations (the method of lines) in two computer labs focusing on different growth terms and boundary conditions. It is possible to include a discussion of medical imaging, such magnetic resonance imaging (MRI) which is clinically used to determine the extent of GBM tumor growth, and its implementation using, for instance, the book by Epstein [35] as an auxiliary resource.
Reaction-diffusion equations also have applications in models of avascular tumor growth and tumor angiogenesis [6, 17, 66, 87]. Discussion of partial differential equations models could lead to students’ interest in advanced techniques and analysis as well as numerical methods.

4. Final Projects and Assessments. As mentioned above, one of our objectives is to incorporate elements of an undergraduate research experience into the course. For this reason, in lieu of a traditional final examination, we require the students to present a 15-minute talk and write a final report on a preselected research paper. A digital or paper poster may be prepared in addition to or instead of the oral presentation. Typically, the paper is one that involves derivation, simulation, and analysis of an ordinary differential equation model of a tumor or related disease process. Table 1 lists some recent papers that we have found to be suitable for this purpose. Alternatively, students may select another paper with approval from the instructors; we strongly prefer papers that allow students to pursue some type of simulation or statistical analysis of data. We have emphasized ordinary differential equation models for simplicity and in view of the students’ prior mathematical preparation. Specific activities that we suggest students might do in their projects are:

- Read the paper fully and be able to give an oral synopsis of the main results.
- Replicate the steps of deriving the mathematical model.
  - Explain the biological and mathematical evidence for the model.
  - Find analytic solutions of the ordinary differential equation model.
  - Determine equilibrium points and their stability.
- Reproduce the main numerical simulations by coding up the model.
  - Reproduce figures showing the results of numerical simulations.

The objective of the project is for students to reproduce some of the main results of the selected paper. We encourage (but do not require) students to work in teams of two. Often, one of the students is better at programming and the other is more confident in front of a classroom audience (as may be the case for students who are not native speakers of English). To assist students in their preparation, we include a lecture on how to give an effective talk, including poise, dress, voice modulation, slide format, and hand gestures. (There are many related resources on the Internet, and Higham’s text [48], which also discusses technical writing, and Alley’s text [2] are useful supplements.) We also require the students to rehearse their talk at least twice with other members of the class.

Students are required to select a paper at least one month before the end of the semester. Towards the end of the term, we set aside one week of class time solely for open office hours, in which the students can work on talks and simulations with input from an instructor. Sometimes, a paper contains a minor error or is not specific about the precise values of model parameters used to generate a figure; in such cases, we suggest workarounds (and sometimes have emailed the authors for clarification).

In the final report, students are asked to include some sample simulations (MATLAB code can be included as an appendix) and to explain in their own words how the model is derived, the main assumptions involved, and the interpretation of the principal numerical results. Usually this can be done in 8–12 pages, and we provide guidance on providing citations and avoiding plagiarism. Beier et al. [10] describe approaches to prepare students for a research/writing experience, examples of specific questions that students can address when working on projects with mathematical models of cancer, and some ideas for assessment.
List of research journal articles that can serve as topics for undergraduate writing/research-based projects organized by type of cancer. Mathematical models in the papers are ordinary differential equation models.

<table>
<thead>
<tr>
<th>Type of Cancer</th>
<th>Papers</th>
</tr>
</thead>
</table>
| General Growth | Benzekry et al. [11]  
Gentry and Jackson [40]  
Ku-Carrillo, Delgadillo, and Chen-Charpentier [60]  
Laird [62, 63]  
Sarapata and de Pillis [89]  
Talkington and Durrett [99] |
| General Treatment | de Pillis, Gu, and Radunskaya [26]  
Feizabadi and Witten [36]  
Jain and Jackson [52]  
Kozusko et al. [59]  
López, Seoane, and Sanjuán [64]  
Murphy, Jaaafari, and Dobrovolny [74]  
Nagy [75]  
Sachs, Hlatky, and Hahnfeldt [88]  
Talkington, Dantoin, and Durrett [98]  
Wilson and Levy [108] |
| Bladder | Bunimovich-Mendrazitskya et al. [15, 16] |
| Brain | Sturrock et al. [93] |
| Breast | Eladdadi and Isaacson [32, 33]  
Enderling et al. [34]  
Miller et al. [70]  
Mufudza, Soroña, and Chiyaka [73]  
Roe-Dale, Isaacson, and Kupferschmid [85] |
| Cervical | Brown and White [14] |
| Colorectal | Bjerknes [12]  
de Pillis, Savage, and Radunskaya [27]  
Johnston et al. [54] |
| Leukemia/Lymphoma | Clapp et al. [19]  
Marciniak-Czochra et al. [68]  
Michor et al. [69]  
Moore and Li [71]  
Nanda, de Pillis, and Radunskaya [76]  
Paquin et al. [82]  
Roesch, Hasenclever, and Scholz [86]  
Stiehl, Lutz, and Marciniak-Czochra [92] |
| Lung | Kang et al. [55]  
Rhodes and Hillen [84] |
| Melanoma | Kogan, Agur, and Elishmereni [57] |
| Pancreatic | Louzoun et al. [65] |
| Prostate | Baez and Kuang [9]  
Gallaher et al. [38]  
Hedican, Kemper, and Lanie [47]  
Jain and Friedman [53]  
Tosoian et al. [105] |
| Renal | dePillis et al. [25] |
5. Discussion. In this paper we present material that can be used to develop an undergraduate course in mathematical modeling of cancer growth and treatment. In particular, general growth models such as exponential, logistic, von Bertalanffy, and Gompertz growth are introduced as well as models of ovarian cancer, prostate cancer, and glioblastoma multiforme. Probability and statistics topics are also covered. Following the content on cancers, resources for effective reading and communication of research are given as well as a list of research journal articles suitable for student projects. In the Supplementary Materials, a sample syllabus, homework assignments, and computer labs are provided. This toolkit will be beneficial to those instructors interested in developing or supplementing a mathematical modeling course.

Student response to the course has been mostly positive; preliminary feedback from students indicates that the course material was found to be very interesting and engaging. The Mathematics + Cancer course has been taught three times: in Fall 2013, 2015, and 2017. Evaluations were lowest the first time the course was taught, as there was a learning process for the instructors as well as the students regarding topics and mathematical level. Among the positive comments in students’ anonymous course evaluations include the novelty of the topics; relevance of mathematics to the “real world;” and the opportunity to give a talk on a paper of interest. Not all students liked the MATLAB exercises, and one (who was not a native English speaker) expressed frustration about some of the medical and biological terminology.

No student has ever complained that the pace of the course was too slow! Our advice to instructors who are interested to offer a course of this nature is to start slowly: a review of the basic concepts of exponential growth and decay, half-life and doubling times, etc., is welcome. It is wise to define biological terms more than once; concepts such as phenotype, signaling, “wild type,” chromosome, receptor, fitness, mutation, etc., can be explained and their implications developed over the course of several lectures. We provide further information and data on student evaluations of the course in the Supplementary Materials.

Acknowledgments. Much of the course content is derived from the authors’ experiences in mentoring undergraduate students as part of the National Science Foundation Mentoring through Critical Transition Points program under grant number DMS-1148771.

The authors would like to thank the anonymous reviewers for their careful reading and many helpful suggestions.

REFERENCES


This manuscript is for review purposes only.
MATHEMATICS + CANCER


This manuscript is for review purposes only.


H. Enderling, M. A. J. Chaplain, A. R. A. Anderson, and J. S. Vaidya, 


H. Jain and T. Jackson, Mathematical modeling of cellular cross-talk between endothelial and

This manuscript is for review purposes only.


[73] C. Mufudza, W. Sorofa, and E. T. Chiyaka, Assessing the effects of estrogen on the


C. Tomasetti and B. Vogelstein, Variation in cancer risk among tissues can be explained by the number of stem cell divisions, Science, 347 (2015), pp. 78–81, http://dx.doi.org/10.1126/science.1260825.


This manuscript is for review purposes only.
SUPPLEMENTARY MATERIALS: MATHEMATICS + CANCER: AN UNDERGRADUATE “BRIDGE” COURSE IN APPLIED MATHEMATICS∗

TRACY L. STEPIEN†, ERIC J. KOSTELICH‡, AND YANG KUANG‡

In these supplementary materials, we provide a sample syllabus, homework problem sets, and computer lab descriptions for use in developing a Mathematics + Cancer course. We also provide information and data on student evaluations of the course.

SM1. Sample Syllabus.

Course Description. This course addresses some of the mathematical questions regarding the dynamics of cancer growth and treatment. We will discuss some classical models of cancer growth and then consider several case studies, which may include prostate cancer, melanoma, chronic myeloid leukemia, and glioblastoma. We will begin with a review of some basic concepts from ordinary differential equations (ODEs), then discuss cancer growth and treatment models that can be formulated as ODEs. Limitations of the models, including an introduction to uncertainty quantification, will also be highlighted. In the latter half of the course, we will study a very important partial differential equation, the diffusion equation, and its connection to models of cancer growth. Our investigations will include analytical, qualitative, and numerical methods, including the use of ODE solvers in MATLAB. During most weeks, one or more research papers will be assigned for reading. Short quizzes on their content will be given at the start of class during the week following the reading assignment. The course will be a lecture-lab format: We will meet in the regular classroom on one day a week for lecture and in the computer lab on the other day. Previous MATLAB programming experience is helpful but not essential. At the end of the course, you will present a 15-minute talk and write an original paper of 8–12 pages in length about a research paper of your choice. The instructor will suggest some suitable papers, but they are happy to consult with you if you wish to suggest another paper that interests you.

Course Outcomes. You will learn how to formulate, solve, parameterize, and validate ordinary differential equation models of certain kinds of cancer and treatment; some of the basic biology of cancer and its treatment; the use and limitations of ode45 and similar solvers in MATLAB; and how to give a talk on mathematical research.

Prerequisites. Three semesters of calculus and one semester of differential equations. Linear algebra is recommended as a pre- or co-requisite.

Assessment. Assessment will be determined from regular written homework assignments, short quizzes based on the content of assigned reading, computer labs, a midterm exam, and a final presentation. During the second half of the semester, you will have the opportunity to work as part of a team (usually 2 people) on a project involving some topic discussed in the course. You will discuss one or two research papers and/or do some numerical simulations of a relevant mathematical model. The

∗This work was supported by the National Science Foundation’s Mentoring through Critical Transition Points program under grant number DMS-1148771.
†School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287-1804. Current address: Department of Mathematics, University of Arizona, Tucson, AZ 85721 (stepien@math.arizona.edu).
‡School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287-1804 (kostelich@asu.edu, kuang@asu.edu).

SM1

This manuscript is for review purposes only.
instructor will suggest some papers, but you are welcome to propose other research topics that interest you. In any case, the final topic and scope of work will be negotiated with the instructors well in advance. In lieu of a final exam, you will give an oral presentation and write a 8–12 page paper (due at the final exam time).

**Tentative Schedule.**

- **Week 1.** Exponential and logistic growth models; direction fields; difference equations.
- **Week 2.** Autonomous differential equations; phase diagrams; the logistic and von Bertalanffy’s growth models.
  (Reading: Hanahan and Weinberg [SM12])
- **Week 3.** Autonomous differential equations; phase diagrams
  (Reading: Couzin [SM6])
- **Week 4.** Mathematical models of chemotherapy and surgery.
  (Reading: Kohandel et al. [SM16])
- **Week 5.** Other formulations of Gompertzian growth; conditional probability.
- **Week 6.** A mathematical model of prostate cancer treatment.
  (Reading: Jannini et al. [SM14], Hirata et al. [SM13], Portz et al. [SM20], Swanson et al. [SM24])
- **Weeks 7–8.** Statistical sampling, hypothesis testing, and uncertainty quantification.
  (Reading: Cohen [SM5], Goodman [SM10])
- **Week 9.** Midterm exam. How to give a scientific presentation. Overview of final talks and papers, including suggested references.
- **Week 10.** The diffusion equation; separation of variables.
  (Reading: Byrne [SM4])
- **Week 11–12.** Glioblastoma biology and treatment; mathematical models of glioblastoma.
  (Reading: Swanson et al. [SM23])
- **Week 13.** Open office and lab hours. Discussion of final talks, papers, and associated computations.
- **Week 14.** Mathematical model of TBA cancer growth.
- **Week 15.** Oral presentations.

**SM2. Homework Problems.** The following sets of problems may be used in conjunction with the content modules for assigned written homework.

**SM2.1. Dynamical Models of General Tumor Growth I.**

1. One breeding pair of cockroaches can produce 1,000 fertile offspring (500 new breeding pairs). If every new breeding pair survives and produces 1,000 offspring, after how many generations would the entire earth be covered 1 meter deep in cockroaches?  (*The surface area of the earth is about $5 \times 10^8$ km$^2$. Assume each cockroach occupies a volume of 1 cm$^3$. You may also assume that the parents die off when the juveniles reach maturity.*)

2. Let $L(t)$ denote the length of a worm at time $t$. von Bertalanffy proposed a growth model for a typical worm given by $L' = r(L_\infty - L(t))$, where $L_\infty$ is the length of a fully grown individual and $r$ is the growth rate. Find a general solution for $L(t)$, where $L_0$ is the initial length of the baby worm.

3. Consider two applications of the separation of variables for ordinary differential equations related to the von Bertalanffy growth model and the Gompertzian growth model:
(a) Show how to derive the solution $S(t) = \frac{a}{\beta} + \left(S_0 - \frac{a}{\beta}\right)e^{-\frac{\alpha}{\beta}t}$ from the ordinary differential equation $\frac{dS}{dt} = \frac{1}{3}(\alpha - \beta S)$.

(b) Show how to derive the solution $L(x) = L_0 \exp\left(\frac{\alpha}{\beta}q(1-q^x)\right)$ from the ordinary differential equation $\frac{dL}{dx} = -\alpha q^x L$.

4. Show how the substitution of the function $x(t) = \ln\left(\frac{K}{N(t)}\right)$ into the differential equation $\frac{dN}{dt} = \alpha N(t) \ln\left(\frac{K}{N(t)}\right)$ produces the exponential decay equation $\frac{dx}{dt} = -\alpha x$.

5. Draw a phase diagram for the differential equation $x' = x(x-1)(2-x)$.

Identify each equilibrium point and state whether it is stable or unstable.

SM2.2. Dynamical Models of General Tumor Growth II.

1. Consider an exponential growth process, $x' = rx$ ($r > 0$). The doubling time is the time $\Delta t$ required such that $x(t + \Delta t) = 2x(t)$. Find $\Delta t$ as a function of $r$ and show that it does not depend on the value of $t$.

2. Consider the general von Bertalanffy model $\frac{dW}{dt} = \alpha W^\lambda - \beta W^\mu$. Show that the origin is a source and $(\alpha/\beta)^{1/(\mu-\lambda)}$ is globally attracting if and only if $\lambda < \mu$. Analyze completely the case in which $\lambda \geq \mu$. Hints:
   - The origin is a source if initial conditions near the origin eventually move away from the origin. To prove that the origin is a source, you need to show that $W' > 0$ if $W(0)$ is suitably small and positive and likewise that $W' < 0$ if $W(0)$ is suitably small and negative.
   - Show by substitution that if $W_{eq} = (\alpha/\beta)^{1/(\mu-\lambda)}$, then $W' = 0$. In other words, an initial population that is exactly equal to $W_{eq}$ remains at a constant size, thus $W_{eq}$ is an equilibrium point.
   - We say that $W_{eq}$ is globally attracting if for every positive choice of $W(0)$, the corresponding solution asymptotes to $W_{eq}$. Demonstrate that $W' < 0$ whenever $W(t) > W_{eq}$ and that $W' > 0$ whenever $W(t) < W_{eq}$.
   - It is helpful to sketch a phase diagram.

3. Let $y = W^{1-\lambda}$. Assuming $W$ satisfies the equation

\begin{equation}
\frac{dW}{dt} = \alpha W^\lambda - \beta W,
\end{equation}

show that $\frac{dy}{dt} = (1-\lambda)(\alpha - \beta y)$. Then show that equation (SM1) with initial condition $W(0) = W_0$ has the solution

$W(t) = \left(\frac{\alpha}{\beta} - \left[\frac{\alpha}{\beta} - W_0^{(1-\lambda)}\right] e^{-(1-\lambda)\beta t}\right)^{\frac{1}{(1-\lambda)}}$, \quad $\lambda \neq 1$.

4. Let

$A = \begin{pmatrix} \frac{1}{2} & \frac{3}{2} \\ \frac{3}{2} & \frac{1}{2} \end{pmatrix}$.

(a) Calculate the eigenvalues and eigenvectors of $A$.

(b) Consider the system $x = Ax$. Is the origin a sink, a source, or a saddle?
SM2.3. Mathematical Models of Chemotherapy and Surgery. The goal of these exercises is to work through the simplest version of the Kohandel et al. [SM16] model of ovarian cancer treatment. In these exercises, we suppose that an ovarian tumor of size \( x(t) \) grows exponentially starting at \( t = 0 \) from an initial seed of size \( X_0 \); that is, \( x' = ax \) for some positive constant \( a \). Also suppose that, when chemotherapy is given, it kills cells at a rate proportional to the current tumor size (the log-kill hypothesis). In other words, the chemotherapy kills cells at the rate of \(-cx(t)\), where \( c \) is another positive constant.

1. What differential equation describes \( dx/dt \) while chemotherapy is given?

What condition on \( a \) and \( c \) must hold for chemotherapy to produce a response (i.e., shrink the tumor)?

2. Consider a situation in which the tumor grows (untreated) from time \( t = 0 \) to time \( t = t_0 \). What is \( x(t_0) \)?

3. Assume that the tumor is diagnosed at \( t_0 \) and treatment begins. Suppose that chemotherapy is given first and continues from \( t = t_0 \) to \( t = t_f \). Find \( x(t_f) \) in terms of \( a \), \( c \), \( t_0 \), \( t_f \), and \( X_0 \).

4. Surgery is performed at \( t = t_f \) and removes \( e^{-k} \) of the tumor cells remaining after chemotherapy. Find \( \lim_{t \to t_f^+} x(t) \), that is, the tumor population immediately after surgery.

5. As in Problem 2, suppose the tumor grows (untreated) from time \( t = 0 \) to time \( t = t_0 \). Now suppose that surgery is performed when the tumor is first diagnosed and removes \( e^{-k} \) of the tumor cells. Find \( \lim_{t \to t_0^+} x(t) \), the cell population immediately after surgery, in terms of \( a \), \( c \), \( t_0 \), and \( X_0 \).

6. Next assume that chemotherapy follows immediately after surgery and continues to \( t = t_f \). Find \( x(t_f) \) in terms of \( a \), \( c \), \( t_0 \), \( t_f \), and \( X_0 \).

7. According to this model, does the sequencing of surgery and chemotherapy affect the final tumor cell population? Justify your answer.

Now assume that the tumor grows according to the logistic growth model \( x' = ax(1 - x/K) \).

8. Show that \( x_{CS} < x_{SC} \), where \( x_{CS} \) and \( x_{SC} \) are as given in the main article in equations ?? and ??, respectively.

SM2.4. Conditional Probability.

1. These exercises consider a way to quantify the effectiveness of drug tests.

(a) Suppose a test for illicit drug use correctly identifies 90% of current users but has a false positive rate of 2.5%. What fraction of people who test positive actually are current users, using data from the 2015 Annual Survey on Drug Use and Health [SM21] that found that about 10.1% of Americans 12 years and older had used an illicit drug in the month prior to the survey (i.e., \( P(D) = 0.101 \))?

(b) Supposing a 90% detection rate as before, what must the false positive rate be to be 95% certain that a person who tests positive uses illicit drugs?

(c) Repeat the previous problem assuming that the test correctly identifies 99% of current users (i.e., \( P(+|D) = 0.99 \)).

2. These exercises explore the likelihood that test results for a rare genetic mutation actually imply that you have the mutation.

(a) Suppose that 1% of the population carries a particular genetic mutation. Also suppose that the test correctly identifies 95% of the people who
carry the mutation and correctly rules out 99% of the population that
does not carry the mutation. What is $P(M | +)$, the probability that you
carry the mutation given that you have tested positive for it?
(b) What is the probability that you do not carry the mutation given that
you test negative?
(c) As in (a), but suppose that 10% of the population carries the mutation.
(d) As in (b), but suppose that 10% of the population carries the mutation.

SM2.5. Uncertainty Quantification. Recall the following two definitions of
error: Suppose the true (nonzero) value of some quantity is $x$ and we measure $\hat{x}$. The
absolute error in the measurement is $|\hat{x} - x|$ and the relative error is
\[
\left| \frac{\hat{x} - x}{x} \right|.
\]
These exercises ask you to estimate errors in solutions of differential equations that
arise from various sources.
1. Consider an exponential growth process, $x' = rx$, where $r > 0$. Suppose that
we have a correct model (i.e., the value of $r$ is known) but that the initial
condition, $x(0) = x_0$, is subject to measurement error. If we measure $\hat{x}_0$
instead, then our predicted value is $\hat{x}(t) = \hat{x}_0 e^{rt}$, but the truth is $x(t) = x_0 e^{rt}$.
Show that the relative error in the prediction at time $t$ is constant.
2. Next suppose that we estimate the growth rate as $\hat{r}$ instead of the true value,
r. Explain what happens to the relative error as a function of time (there are
two cases).
3. Finally, consider a logistic growth model, $x' = rx(1 - x/K)$, where there are
two parameters, the growth rate $r$ and the carrying capacity, $K$. You may
use an analytical or graphical argument to justify each of your answers.
(a) If $r$ and $K$ are known but the initial condition $\hat{x}_0$ is subject to measure-
ment error, what do you expect will happen to the relative error in any
prediction as $t \to \infty$?
(b) As in (a), but suppose that the growth rate $\hat{r}$ is subject to error.
(c) As in (a), but suppose that the carrying capacity $\hat{K}$ is subject to error.

1. Consider the ODE
\[
\frac{d^2 \phi}{dx^2} + \lambda \phi = 0.
\]
For each set of boundary conditions in the list below, determine the conditions
that $\lambda$ must satisfy so that there is a unique and real-valued solution. You
may assume that $\lambda > 0$ in your analysis.
(a) $\phi(0) = 0$ and $\phi(\pi) = 0$
(b) $\phi(0) = 0$ and $\phi(1) = 0$
(c) $\phi(0) = 0$ and $\frac{d\phi}{dx}(L) = 0$
(d) $\frac{d\phi}{dx}(0) = 0$ and $\phi(L) = 0$
(e) $\frac{d\phi}{dx}(0) = 0$ and $\frac{d\phi}{dx}(L) = 0$
2. Consider the heat equation on a rod of length $L$,
\[
\frac{\partial u}{\partial t} = k \frac{\partial^2 u}{\partial x^2},
\]
with the boundary conditions \( u(0, t) = u(L, t) = 0 \). For each initial condition in the list below, find the solution \( u(x, t) \).

(a) \( u(x, 0) = 6 \sin \left( \frac{2\pi x}{L} \right) \)

(b) \( u(x, 0) = 3 \sin \left( \frac{\pi x}{L} \right) - \sin \left( \frac{3\pi x}{L} \right) \)

3. Show how to apply the separation of variables procedure to solve the heat equation

\[
\frac{\partial u}{\partial t} = k \frac{\partial^2 u}{\partial x^2}
\]

subject to the boundary conditions

\[
\frac{\partial u}{\partial x}(0, t) = \frac{\partial u}{\partial x}(L, t) = 0.
\]

These no-flux boundary conditions imply that there is no flow of heat across the ends of the rod (i.e., the ends are perfectly insulated). Find a general solution of the heat equation \( u_n \).

**SM3. Computer Labs.** The following sets of computer labs may be used in conjunction with the content modules. All of the labs are designed for MATLAB [SM25], but the labs can be implemented in other mathematical computing programs such as GNU Octave [SM9], Maple [SM17], Mathematica [SM26], and CPython [SM7].

**SM3.1. MATLAB Preliminaries.** In the case where students need to be (re)oriented with MATLAB, it is suggested that the first computer lab be an introduction. Since students will write a lot of function files throughout the course, it is important to develop an understanding of entering commands, writing conditionals and loops, plotting, and creating script files. Many materials are available online that can serve as a basis for an introductory lab, for example, through the MathWorks web site [SM25].

**SM3.2. Solving First-Order Ordinary Differential Equations.** To introduce students to function files and solving first-order ordinary differential equations with \texttt{ode45} in MATLAB, we focus a lab on the exponential equation, logistic equation, and the Gompertz equation which are all of the form

\[
\frac{dy}{dt} = f(t, y).
\]

First, create function files for the right hand sides of the equations \( f(t, y) \), where in particular we have

(a) \texttt{(SM3a)} exponential: \( f(t, y) = \alpha y \),

(b) \texttt{(SM3b)} logistic: \( f(t, y) = \alpha y \left( 1 - \frac{y}{\beta} \right) \),

(c) \texttt{(SM3c)} Gompertz: \( f(t, y) = \alpha y \ln \left( \frac{\beta}{y} \right) \),

with \( \alpha \) and \( \beta \) as parameters. For some of the functions, choose exact values for \( \alpha \) and \( \beta \), but for at least one of the functions, let the parameters be unspecified so that they are inputs into the function file.
Once the function files are written, the ordinary differential equation (SM2) can be numerically solved with \texttt{ode45} given an initial condition and interval of integration. Some variations that can be explored are:

- plot both the numerical and true solutions;
- solve with various initial conditions and plot the results; and
- given sample data and the exponential equation, find an initial condition and parameter value $\alpha$ for which the solution best fits the data using guess and check (limit possible values to integers).

**SM3.3. Analyzing and Solving Systems of First-Order Ordinary Differential Equations.** After examining single equations, another lab can focus on systems of first-order ordinary differential equations. A good system of equations that can be studied is the Gyllenberg–Webb model [SM11]. Gyllenberg and Webb observed that actively proliferating cells in tumors can enter a quiescent state when they stop dividing and quiescence tends to be more common in large tumors compared to small tumors. They developed a mathematical model of the transition of cells from a proliferative state to a quiescent state and back. The model equations are

\[
\begin{align*}
\frac{dP}{dt} &= (\beta - \mu_p - r_0(N)) P + r_i(N) Q, \\
\frac{dQ}{dt} &= r_0(N) P - (r_i(N) + \mu_q) Q,
\end{align*}
\]

where $P(t)$ is the number of proliferative cells, $Q(t)$ is the number of quiescent cells, and $N(t) = P(t) + Q(t)$ is the total number of cells. The parameters are the proliferation rate $\beta > 0$, death rate for proliferative cells $\mu_p \geq 0$, and death rate for quiescent cells $\mu_q \geq 0$. The proliferative cells transition to and from being quiescent by rates $r_0(N)$ and $r_i(N)$, respectively. We let the transition rates be

\[
\begin{align*}
r_0(N) &= \frac{kN}{aN+1}, \\
r_i(N) &= \frac{r}{N+m},
\end{align*}
\]

and note that $k$, $a$, $r$, and $m$ are parameters.

It is useful to determine what effects the parameters and initial conditions $P(0)$ and $Q(0)$ have on the solutions, and this can be analyzed by investigating phase portraits (for example with pplane [SM19]) and then numerically solving the system of differential equations (SM4). Parameter values can be doubled and halved one at a time and the solutions can be plotted to compare and contrast the results.

**SM3.4. Mathematical Models of Chemotherapy and Surgery.** After introducing the model of chemotherapy and surgery treatment sequencing in ovarian cancer of Kohandel et al. [SM16] in lecture, students can numerically simulate and analyze the different results when comparing adjuvant versus neoadjuvant chemotherapy and the three cell-kill hypotheses. One could also look at the differences between the different tumor growth model functions $f(x)$, but it may be more beneficial to choose just one. Since the students will have already written function files for the exponential, logistic, and Gompertz right hand side functions in the lab described in Subsection SM3.2, they should be able to implement them into their code for this lab.

The next step is for students to write function files for the right hand side of equations ?? and ?? for each of the log-kill hypotheses given in equations ??????. Then for each of the log-kill hypotheses, a script file can be written to examine adjuvant chemotherapy and neoadjuvant chemotherapy.
For adjuvant therapy, equation \( t_f \) is solved first on the interval of integration \([0, t_f]\). Then the last entry of the \( x \) solution vector (multiplied by \( e^{-k} \)) is used in the initial condition to solve equation \( t_f \) on the interval of integration \([t_s, t_f]\). The final tumor size \( s \) is just the last entry of the \( x_c \) solution vector. The overall solution vectors from the initial value problems \( t_f \) can be combined into one longer vector for plotting.

For neoadjuvant therapy, the only change that needs to be made is within the initial conditions for the second half of the simulation. After the first part \( t_f \) is solved, the initial condition for the second equation \( t_f \) is just the last entry of the \( x_c \) solution vector. However, the final tumor size \( s \) uses the last entry of the \( x_c \) solution vector, but it needs to be multiplied by \( e^{-k} \).

To compare the different simulations, students can

- plot the adjuvant and neoadjuvant solutions on the same graph for each of the cell-kill hypotheses;
- determine if the order of sequencing of surgery and chemotherapy affects the final number of tumor cells; and
- determine which of the three cell-kill hypotheses results in the smallest tumors by time \( t_f \).

For each of the situations where neither the adjuvant or neoadjuvant chemotherapy treatment methods resulted in a small tumor size compared to the initial size, students can find parameters (out of \( c, k, \) and \( \delta \)) that result in a larger increase in tumor size.

### SM3.5. Conditional Probability

An example of conditional probability occurs in testing a drug’s efficacy in treating a disease. In this lab, we simulate a clinical trial for a hypothetical cancer drug designed to reduce tumor size. This lab was based off of the “Biological Example of Conditional Probability: Drug Testing” section and DrugTesting.m MATLAB code in Bodine et al. [SM3].

The hypothetical clinical trial has 200 patients, and half receive the cancer drug and half receive a placebo. The number of patients who received the cancer drug and experienced a decrease in tumor size versus increase and the number of patients who received the placebo and experienced a decrease in tumor size versus increase are given to the students. Ultimately we would like to know: what is the probability that if you take the cancer drug \( A \) then the size of your tumor will decrease \( B \), i.e., what is \( P(B|A) \)? The true value of \( P(B|A) = P(A \cap B)/P(A) \) can be calculated to compare to results from the lab.

To simulate the clinical trial, \( N \) experiments are run \( N \) times, where \( N \) is the number of patients in the clinical trial, in which we randomly pick one patient and record if they were given the cancer drug or placebo and if the size of their tumor increased or decreased. The probability of “took cancer drug” \( A \) and the probability of “took cancer drug” and “tumor decrease” \( A \cap B \) for each set of experiments is added to a running sum. These summed values approximate \( P(A) \) and \( P(A \cap B) \), which can be divided to approximate \( P(B|A) \).

Using the estimate of \( P(B|A) \), the lab can be extended to determine \( P(A|B) \) with Bayes’ Theorem \( \Box \). While Bayes’ Theorem does not present the quickest way to code the probabilities, it allows the students to gain more experience with conditional statements (if, then, else) by estimating \( P(B), P(A^C), \) and \( P(A^C \cap B) \).

### SM3.6. Statistical Sampling

Suppose we want to know the average height of a group of individuals, but we are unable to measure the height of every single person in the group. How many people do we need to measure to get a reasonable estimate?

To explore sample means, we use a data set that lists the height in a sample of 1794
pregnant women (Bland [SM2]; data available for download at http://www-users.york.ac.uk/~mb55/datasets/datasets.htm as “Height”). We will explore how the mean, median, variance, and standard deviation changes based on the sample size of the data, and plot histograms and box plots to help visualize what is going on. First, besides having the full data set, the data should be also grouped into smaller samples of various sizes (e.g., 10, 20, 40, 80, 120, 320, 640, and 1280). This can be done ahead of time so that there is uniformity between the students’ results, or students can create the data subsets by randomly selecting the data points.

For each sample set, have students calculate the mean, median, variance, and standard deviation as well as plot a histogram with the normal distribution probability density function overlaid on top. In one figure, plot the box plot for all of the data sets. With this information, students can infer what the minimum sample size is that gives a reasonable estimate of the average height.

SM3.7. Hypothesis Testing. A result is called statistically significant if it has been predicted as unlikely to have occurred by chance alone, according to a pre-determined threshold probability, the significance level. We study survival data for advanced lung cancer patients (Kalbfleisch and Prentice [SM15]; data available for download at ftp://ftp.wiley.com/public/sci tech_med/failure_time/ as “Data Set I” or at http://lib.stat.cmu.edu/DASL/). The main purpose of the study was to compare the effects of two chemotherapy treatments in prolonging survival time, termed “standard” and “test.” The patients can have different types of tumors and they have been classified into four categories: squamous, small, adeno, and large. We compare the chemotherapy treatments for the four tumor size categories separately.

The null hypothesis to test with this data set is that there is no difference in average survival time between receiving “standard” treatment or “test” treatment, i.e. \( H_0 : \mu_{test} = \mu_{standard} \), while the alternative hypothesis is that there is a difference in average survival time between receiving “standard” treatment or “test” treatment, i.e. \( H_a : \mu_{test} \neq \mu_{standard} \). We choose to test at a significance level \( \alpha = 0.05 \).

The statistic of interest is the \( t \)-statistic

\[
(t) = \frac{\bar{x}_{test} - \bar{x}_{standard}}{\sqrt{\frac{s^2_{test}}{n_{test}} + \frac{s^2_{standard}}{n_{standard}}}},
\]

where \( \bar{x}_T \) is sample mean, \( s_T \) is the sample standard deviation, and \( n_T \) is the sample size (number of individuals) for groups \( T = \text{“test”} \) and \( \text{“standard”} \).

After students count the number of patients with squamous, small, adeno, and large tumors with survival \( \leq 100 \) days, \( 400 \) days < survival \( \leq 300 \) days, and survival > 300 days (12 categories total), then sample means, sample standard deviations, and sample sizes can be calculated. Then the \( t \)-statistic (SM6) can be calculated to determine whether to reject or fail to reject the null hypothesis. Students can comment on for which size tumors there was a statistically significant difference in average survival time between the “standard” and “test” treatments and explain what it means in terms of what the results of the clinical trial imply about the effectiveness of the treatments for the different size tumors. Then students can vary \( \alpha \) to find cases where more of the tests results in rejecting the null hypothesis.

SM3.8. Simulating a Prostate Cancer Growth Model. The goal of this lab is to reproduce the figures in Portz et al. [SM20], which models the effects of androgen-deprivation therapy in prostate cancer and compares numerical simulations with actual patient data. This lab can be split into 2 or 3 sessions.
SM3.8.1. Model Equations. The model of Portz et al. [SM20] is motivated by the Droop [SM8] model as discussed in ???. Population dynamics of androgen-dependent (AD) cells and androgen-independent (AI) cells are described by the following equations,

\[
\begin{align*}
\frac{dX_1}{dt} &= \mu_m \left( 1 - \frac{q_{\text{min},1}}{Q_1} \right) X_1 - \frac{d_1 X_1}{\text{cell quota}} - \frac{\lambda_1(Q_1) X_1}{\text{mutations}} + \frac{\lambda_2(Q_2) X_2}{\text{mutations}}, \\
\frac{dX_2}{dt} &= \mu_m \left( 1 - \frac{q_{\text{min},2}}{Q_2} \right) X_2 - \frac{d_2 X_2}{\text{cell quota}} - \frac{\lambda_2(Q_2) X_2}{\text{mutations}} + \frac{\lambda_1(Q_1) X_1}{\text{mutations}},
\end{align*}
\]

where \(X_1\) is the number of AD cells, \(X_2\) is the number of AI cells, androgen is the cell quota \(Q_i\) in cells \(X_i\) for \(i = 1, 2\), \(d_1\) and \(d_2\) are the constant rates of cell death, \(\lambda_1(Q_1)\) is the transition rate from AD to AI due to mutations, and \(\lambda_2(Q_2)\) is the transition rate from AI to AD. AD cells grow faster as \(Q(t)\) increases and while AI cells need androgen, they can grow at much lower levels (hence \(q_{\text{min},1} > q_{\text{min},2}\)).

It is assumed that the transition rates due to mutations are of the form of Hill functions,

\[
\begin{align*}
\lambda_1(Q) &= c_1 \left( \frac{K_1^3}{K_1^3 + Q^3} \right), \\
\lambda_2(Q) &= c_2 \left( \frac{Q^3}{K_2^3 + Q^3} \right),
\end{align*}
\]

where \(c_1, c_2, K_1,\) and \(K_2\) are parameters. With these assumptions, the AD→AI transition rate yields a low mutation rate for normal androgen levels and a high rate for low androgen levels, while the AI→AD transition rate is high for normal androgen levels and low for low levels.

The Portz et al. [SM20] model also includes prostate-specific antigen (PSA), as this is a quantity that can be clinically measured, unlike the amount of AD and AI cells. It is assumed that PSA is produced at a rate proportional to the cell populations \(X_1\) and \(X_2\) and cell quotas \(Q_1\) and \(Q_2\) and is cleared from the blood at a constant rate, thus

\[
\frac{dP}{dt} = \sigma_0(X_1 + X_2) + X_1 \left( \frac{\sigma_1 Q_1^4}{Q_1^4 + \rho_1^4} \right) + X_2 \left( \frac{\sigma_2 Q_2^4}{Q_2^4 + \rho_2^4} \right) - \delta P,
\]

where \(P\) is PSA, \(\delta\) is the removal rate, and \(\sigma_0, \sigma_1, \sigma_2, \rho_1,\) and \(\rho_2\) are parameters.

Furthermore, it is assumed that the cell quota \(Q_i(t)\) changes over time for each cell population and it depends on the androgen suppression therapy given to each patient. The derivative \(Q_i(t)\) depends on the cell quota, drug dosage, and androgen consumption with the cells, and the equations for \(Q_i(t)\) are

\[
\frac{dQ_i}{dt} = v_m \left( \frac{q_m - Q_i}{q_m - q_{\text{min},i}} \right) \left( \frac{A}{A + v_h} \right) - \mu_m(Q_i - q_{\text{min},i}) - bQ_i,
\]

where \(v_m, q_m, A, v_h,\) and \(b\) are parameters that are assumed to be the same for each cell population.
SM3.8.2. Simulations. The first step is to obtain the prostate-specific antigen (PSA) data for Case #1–7 (Figures 4–10, solid dots, “serum PSA”) and the androgen data (open dots, “serum testosterone”) from Akakura et al. [SM1]. This can be done for the students or it could be part of the lab. An application such as Plot Digitizer [SM18] can be used. Note that the data in Akakura et al. [SM1] is given in months while the simulations in Portz et al. [SM20] are done in days, so the extracted data should be converted to days.

The androgen data $A(t)$ appears in the cell quota $Q_1$ and $Q_2$ equations, and the data needs to be interpolated for use in numerical simulations. One method is to use an exponential fit as given in Equation (13) in Portz et al. [SM20], however, it is suggested to make use of the pchip command in MATLAB for purposes of this lab.

The equations to be simulated are (SM7)–(SM10). In particular, the variables are $X_1$, $X_2$, $Q_1$, $Q_2$, and $P$, and the parameters are $\mu_m$, $q_{\text{min},1}$, $q_{\text{min},2}$, $d_1$, $d_2$, $c_1$, $c_2$, $K_1$, $K_2$, $v_m$, $v_h$, $b$, $\sigma_0$, $\sigma_1$, $\sigma_2$, $\rho_1$, $\rho_2$, and $\delta$. For all 7 cases, the parameter values

$$c_1 = 0.00016, \quad c_2 = 0.00012, \quad K_1 = 0.8, \quad K_2 = 1.7, \quad q_m = 5,$$

$$v_m = 0.27, \quad v_h = 4, \quad b = 0.09, \quad \rho_1 = 1.3, \quad \rho_2 = 1.1, \quad \delta = 0.08,$$

(SM11) can be used, as given in Table I in Portz et al. [SM20], and the remaining parameter values that can be used are given in Table SM1.

Each case should be simulated for the time in days that there is data. In terms of initial conditions, for the cell quotas the following can be used for all 7 cases,

$$Q_1(0) = Q_2(0) = 0.4.$$

For the initial PSA concentration $P(0)$, the MATLAB command pchip can be used to extrapolate the PSA data at $t = 0$. Lastly, it is assumed that the PSA concentration $P$ is a linear combination of cell concentrations $X_1$ and $X_2$, thus we can write this as

$$X_1(0) = \alpha \beta P(0), \quad X_2(0) = (1 - \alpha) \beta P(0),$$

(SM13) for some $\alpha$ and $\beta$, which for all 7 cases let $\alpha = 14/15$, while $\beta = 1$ for Cases #1-4, $\beta = 0.6$ for Case #5, $\beta = 0.7$ for Case #6, and $\beta = 0.8$ for Case #7.

At this point, simulations can be run to reproduce Figures 6–7 in Portz et al. [SM20]. It is suggested that students attempt to reproduce the figures such that all colors, line styles, axes, and legends are the same. Since the parameter values given in Table SM1 and (SM11) are truncated from the exact values used in Portz et
al. [SM20], note that the reproduced curves will be not be exactly the same as the
originals.

The last topic that can be covered in this lab is error estimates, in particular the
error in the PSA concentration $P$ found in the numerical simulation compared to the
data from Akakura et al. [SM1]. After interpolating the numerical solution for $P$,
students can calculate absolute error

$$\|\hat{x} - x\|,$$

relative error

$$\left\| \frac{\hat{x} - x}{x} \right\|,$$

and mean square error

$$\frac{1}{N} \|\hat{x} - x\|^2,$$

where $x$ is a vector of values from data, $\hat{x}$ is a vector of values from simulations, and
$N$ is the length of the vectors. The mean square error can be compared to the values
obtained in Table II in Portz et al. [SM20].

Further questions to discuss regarding the error include: does looking at figures
or the absolute, relative, or mean square error give the best representation of the error
for the simulations in this lab? Does the case with the smallest calculated error match
with which figure appears to match the data the best?

**SM3.9. Solving the Heat Equation.** While there are many examples of how
to use MATLAB to solve partial differential equations like the heat equation, we
explore one aspect of the numerical details using a finite differences scheme. In this
lab, the heat equation on the interval $[0, L]$ will be converted into a system of ordinary
differential equations, which will be solved using `ode45`. Before coding anything in
MATLAB, students should be instructed to write down the details for the derivation
of the numerical ordinary differential equations. The outline of the scheme is:

1. Subdivide interval $[0, L]$ into $n$ equal subintervals; noting that the resulting
   grid has $n+1$ points. Let $\Delta x$ be the length of each subinterval, and let $x_1 = 0$
   be the leftmost grid point and $x_{n+1} = L$ be the rightmost grid point.

2. For a given fixed time $t$, let $u_i = u(x_i, t)$. Approximate $\partial u/\partial x$ at $x_i$ by the
   forward finite difference

$$\Delta u_i = \frac{u_{i+1} - u_i}{\Delta x}, \quad i = 1, 2, \ldots, n,$$

and approximate $\partial^2 u/\partial x^2$ by the finite difference

$$\Delta^2 u_i = \frac{\Delta u_i - \Delta u_{i-1}}{\Delta x}, \quad i = 2, 3, \ldots, n.$$

Expand the formula on the right-hand side of equation (SM18) using equation (SM17)
to find an expression for $\Delta^2 u_i$ in terms of $u_{i+1}$, $u_i$, and $u_{i-1}$.

3. Assuming the Dirichlet boundary conditions $u(0, t) = u(L, t) = 0$, the heat
equation is approximated by the system of ordinary differential equations,

$$u_i' = k \Delta^2 u_i, \quad i = 2, 3, \ldots, n,$$

where the prime denotes the time derivative, and $u_1 = u_{n+1} = 0.$
To analyze the results, students can:

- compare the numerical solution to the true solution;
- change the number of subintervals \( n \); and
- plot the solution versus \( x \) and determine what happens over time.

**SM3.10. Simulating a Model of Brain Tumor Growth.** The brain tumor growth model of Swanson et al. [SM22, SM23] is described by a reaction–diffusion equation. In one dimension on the spatial domain \([0, L]\), this equation is

\[
\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2} + \rho u(1 - u), \quad 0 \leq x \leq L, \quad t > 0,
\]

where \( u(x, t) \) is the concentration of tumor cells (as a percentage of the carrying capacity). \( D \) is the diffusion constant that gives a measure of undirected (random) motion of the cells and \( \rho \) is the proliferation rate.

Considering brain tumor growth in one dimension, a biologically reasonable domain is a nerve fiber. In this case, we will assume the two ends of the nerve fiber are insulated, which corresponds to no-flux boundary conditions, i.e., the Neumann conditions

\[
\frac{\partial u(0, t)}{\partial x} = 0, \quad \frac{\partial u(L, t)}{\partial x} = 0, \quad t \geq 0.
\]

We also assume that a certain concentration of tumor cells exists at the “initial” time, which would correspond to the time at diagnosis, for example

\[
u(x, 0) = \begin{cases} 
\frac{1}{2}, & \frac{L}{3} \leq x \leq \frac{2L}{3} \\
0, & \text{otherwise}
\end{cases}.
\]

The two main differences between the lab described in Subsection SM3.9 are the proliferation term \( \rho u(1 - u) \) and the boundary conditions. The proliferation term is approximated by adding \( \rho u_i(1 - u_i) \) to the discretization (SM19).

For the boundary conditions, we use ghost points \( u_0 \) and \( u_{n+2} \). Discretize the boundary conditions (SM21) with the centered differences

\[
\Delta u_i = \frac{u_{i+1} - u_{i-1}}{2\Delta x}, \quad i = 1,
\]

\[
\Delta u_n = \frac{u_{n+1} - u_{n-1}}{2\Delta x}, \quad i = n + 1.
\]

Then setting (SM23a) and (SM23b) equal to zero, substitute \( u_0 \) and \( u_{n+1} \) into the discretizations for \( u'_{1} \) and \( u'_{n+1} \).

Students can visualize the results of the simulation by plotting the solution at 5 equally spaced times. Parameters can be varied to see how smaller or larger diffusion or proliferation rates can affect the solution.

**SM4. Student Evaluations of Course.** The Mathematics and Cancer course has been taught three times: in Fall 2013, 2015, and 2017. Table SM2 summarizes some of the student responses on a standardized course evaluation form. Answers are reported as the mean value on a 5-point Likert scale (5=strongly agree, 1=strongly disagree). Approximately 40 percent of the students have been women, and, on average, about 10 percent of the initial cohort of students have dropped the course by mid-semester. Overall, 40 to 50 percent of each student cohort have been mathematical sciences majors; 12 to 20 percent, engineering majors; and the remainder from chemistry, life sciences, economics, and other majors.
Table SM2

<table>
<thead>
<tr>
<th></th>
<th>Fall ’13</th>
<th>Fall ’15</th>
<th>Fall ’17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final enrollment</td>
<td>17</td>
<td>16</td>
<td>25</td>
</tr>
<tr>
<td>Number of survey respondents</td>
<td>6</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Assignments/exams promote learning</td>
<td>4.2</td>
<td>4.7</td>
<td>4.4</td>
</tr>
<tr>
<td>Class sessions are well planned</td>
<td>3.5</td>
<td>5.0</td>
<td>4.1</td>
</tr>
<tr>
<td>Pace/level are appropriate</td>
<td>2.8</td>
<td>4.8</td>
<td>4.5</td>
</tr>
<tr>
<td>Course engages my interest</td>
<td>4.0</td>
<td>4.7</td>
<td>4.4</td>
</tr>
<tr>
<td>I learned something valuable</td>
<td>4.2</td>
<td>4.7</td>
<td>4.5</td>
</tr>
</tbody>
</table>

REFERENCES


