21

Predictive Mathematical Modeling in Metastasis

Jonathan A. Sherratt

1. Introduction

Mathematical modeling is emerging as a powerful predictive tool in many areas of biology and medicine, with applications to cancer metastasis increasingly widespread and effective. This type of modeling involves quantitatively accurate representations of specific cellular activities, and is quite different from more traditional applications of mathematics to cancer, such as the simple fitting of experimental data to Gompertzian growth curves. It is made possible by the twin revolutions in molecular biology and nonlinear mathematics over the last two decades, and involves using experimental data at the cell and molecular level to construct mathematical models, which can then be used to predict the macroscopic implications of this data.

Mathematical models have been in use for biological prediction since the early part of the century, initially in ecology and embryology. These early models were phenomenological, that is, they acted as a convenient way to express and explore theories, but did not represent particular postulated mechanisms. Establishment of mathematical biology as a recognized scientific field was achieved by a number of major successes for these early models. Most notable amongst these is the work of Hodgkin and Huxley on electrical signaling in nerve axons, which underlies much of neurophysiology, and for which they were awarded the Nobel Prize for Physiology and Medicine in 1963. More recently, the ability to isolate biological mechanisms at the molecular level has led to a new type of mathematical model, which represents specific low-level mechanisms, either known or hypothesised. This use of mathematical modeling was pioneered during the late 1970s and early 1980s by James Murray at the University of Oxford, mainly in applications to developmental biology. An

From: *Methods in Molecular Medicine, vol. 57:*
Metastasis Research Protocols, Vol. 1: Analysis of Cells and Tissues
Edited by: S. A. Brooks and U. Schumacher

example of a success story from this work is the determination of the mechanism of aggregation in the starvation response of cellular slime molds. These normally live as individual cells, but in starvation conditions the individual cells aggregate to form a multicellular slug, several millimeters long, which crawls in a coordinated manner and then redistributes the individual cells in a new location. Mathematical models showed that chemicals secreted by the cells when starved form spiral signaling waves that cause the initial cellular aggregation *(1)*.

The ability that now exists to mathematically model specific cellular activities in medically important contexts is a direct progression from this earlier work. Within cancer biology, the range of potential applications is very wide. The tumor–immune system interaction has been extensively modeled, addressing the role of the immune system in tumor composition and morphology, and the potential for different immunotherapeutic strategies *(2,3)*. The growth constraints on tumors prior to vascularization have also been modelled widely, leading to predicted relationships between cellular parameters and size limitation *(4,5)*. In the area of metastasis, angiogenesis and invasion have both been studied using mathematical modeling. In the former case, work initially focussed on the basic mechanisms via which angiogenic factors stimulate vascular ingrowth *(6)*, and recent achievements include prediction of the ways in which the extracellular matrix can modulate the cellular response to angiogenic factors *(7)*, and simulation of antiangiogenesis strategies *(8)*. Cancer invasion is an ideal topic for mathematical modeling because of the recent discovery of many underlying molecular mechanisms: Models are able to predict the larger scale implications of these mechanisms. This approach has been used to study the delicate balance between proteolytic enzymes and their inhibitors *(9)* and the role of pH gradients in the invasive process *(10)*.

2. Materials

2.1. Computer Hardware

The basic materials requirement is an appropriate computer; PC, Apple Macintosh and Unix Workstation frameworks are all suitable. Many modeling studies do not require levels of processing power that are problematic for most modern machines. However, investigation of behavior in two or three space dimensions does often demand high level computing, in terms of processing power, RAM, and graphics. Before embarking on such a study, it is advisable to model the behavior concerned in one spatial dimension.

2.2. Computer Software

Some types of mathematical model can easily be simulated directly using commercially available computer packages such as Maple *(11)* or Mathematica *(12)*; this includes models involving variation in time but not in space, and

some discrete-cell models such as cellular automata. Other types of model require preprocessing before they can easily be simulated: This is particularly true for the "continuum models" discussed in **Subheading 3.2.** For those familiar with the computer languages Fortran or C, writing source code is often the most efficient method of simulation in these cases, as there are a number of relevant subroutines available free of charge, notably from the Higher Education National Software Archive (ftp://unix.hensa.ac.uk/mirrors/netlib/). Another useful software source is the Clinical and Biomedical Computing Unit at the University of Cambridge, UK, which has produced a number of software packages specifically designed for medical problems (*see* http://www.cbcu.cam.ac.uk/software.htm for details). If none of these approaches is suitable, those unfamiliar with numerical simulation are best advised to consult a local applied mathematician.

2.3. Mathematical Expertise

In practice, the majority of mathematical modeling studies occur as a collaboration between an experimental or clinical scientist and a mathematician. A major advantage of this approach is that it enables traditional "pen and paper" analysis of the model. This can often give significant insights into the underlying biology, yielding formulae for predicted behavior as a function of cell and molecular parameters. However, the basic processes of model development and simulation do not require high levels of mathematical expertise.

3. Methods

3.1. Suitability of Mathematical Modeling

The essential first step in developing a mathematical model is to identify a problem that is suitable for a theoretical approach. In general terms, such problems fall into one of two categories: (1) Quantitative predictions, for which the key criterion is that reliable quantitative data exists for all the model inputs. Here the objective of the work will be to predict a macroscopic feature of the system, such as the amount of an anti-angiogenic factor required to reduce the angiogenic response by a given percentage. (2) Qualitative hypothesis testing. It is a common misconception that mathematical modeling is effective only for quantitative prediction. In fact, some of the most effective modeling in cancer biology has been of a qualitative nature. In this case, the objective is to test the feasibility of one or more qualitative hypotheses. For example, it is known that in breast carcinomas, tumor-associated macrophages, which are an important regulator of angiogenesis, are found in clusters distinct from vascular hot spots. One hypothesis is that the macrophage clustering might arise prior to vascularization; a model can be used to test this, determining the precise conditions under which prevascularization patterning arises. (For details of this example, *see* **refs.** *3* and *13*).

It is important to stress that the objective of mathematical modeling work is not to generate a large-scale computer simulation of the whole metastatic cascade; although such a model is feasible in principle, its complexity would make it so sensitive to underlying assumptions as to be of no practical value. Rather, models are used in a very focused way, to study in detail particular aspects of metastasis, which may be as limited as a single cell–cytokine interaction, for example. Thus the philosophy of using a mathematical model is similar to that underlying the use of an in vitro experiment to help in the understanding of an in vivo system.

3.2. Choice of Modeling Framework

The most fundamental choice when developing a mathematical model is the choice of mathematical framework used to represent the biological problem. The approach with the longest history is continuum modeling, in which the discreteness of individual cells is neglected, and local cell densities are used as model variables. This is very well established, and gives mathematical equations closest to those used in more traditional areas of applied mathematics (such as fluid dynamics and solid mechanics). A number of good reference books are available on models of this type in biology and medicine, for example *(14,15)*.

The alternative to a continuum model is the representation of cells as discrete objects. Discrete models are very well established in some other areas of biology, most notably ecology; the articles *(16,17)* compare various different discrete and continuous models for the same ecological phenomena, and are useful background reading when choosing a model formulation. However, in cell biology, discrete models are relatively new and there is not yet a coordinated body of expertise. Successful approaches include direct computational models that track cell boundaries *(18)*, cellular automata *(19)*, and discrete cells studied in a continuous extracellular matrix *(20)*. There is currently no reference book describing these different approaches, and thus the choice of an appropriate discrete formulation requires the reading of original research papers, making the modeling procedure rather more involved. The great advantage of discrete models is their ability to predict phenomena at the individual cell level, for example, the paths taken by individual cells in an invasion assay.

3.3. Procedure for Model Development

1. Selection of variables is a key step in the modeling process. Variables can include cell types, chemical regulators, extracellular matrix components, and cell surface receptors. The most effective models usually have only a small of variables, because this makes it much easier to draw clear precise conclusions. However, in some cases it is helpful to begin developing a larger model as an aid to the derivation of an appropriate simpler form.

than a change in numerical parameters, such as a reduction in the time step. Results that are biologically reasonable but whose precise form is sensitive to numerical parameters usually correspond to a numerical scheme that is stable but has not yet converged to the actual solution. A useful rule of thumb is that time/space steps should be reduced until they no longer visibly affect the solution; detailed numerical tests for convergence are described in numerical analysis textbooks, for example *(21)*.

2. Refinement of the model in the light of initial simulations is very common. For instance, the results may indicate that one particular term in the model is central to the predictions, in which case it may be appropriate to include a more detailed representation of this term. An example of this is provided by work on cell chemotaxis. This is conventially modeled as a movement of cells up gradients of chemoattractant, at a rate dependent on the local chemoattractant concentration. However, in some cases, the predicted behavior depends crucially on the details of this concentration dependence, and in such cases more detailed representations are used in which cell surface receptors are added as explicit variables in the model *(22)*.
3. Reformulation of the model is occasionally appropriate; this means alteration of a basic aspect of the model, rather than simply addition of a new feature as in **Note 2**. Reformulation is required when the model predictions are inconsistent with a key modeling assumption. For example, if a continuum model predicts a spatial pattern on the scale of one or two cell diameters, then the prediction will not be reliable and should be checked using a discrete model.

References

1. Tyson, J. J. and Murray, J. D. (1989) Cyclic AMP waves during aggregation of *Dictyostelium* amoebae. *Development* **106,** 421–426.
2. Adam, J. A. and Bellomo, N. (1997) *A Survey of Models for Tumor-Immune System Dynamics*. Birkhäuser, Boston.
3. Owen, M. R. and Sherratt, J. A. (1997) Pattern formation and spatiotemporal irregularity in a model for macrophage-tumor interaction. *J. Theor. Biol.* **189,** 63–80.
4. Ward, J. P. and King, J. R. (1997) Mathematical modeling of avascular tumor growth. *IMA J. Math. Appl. Med. Biol.* **14,** 39–69.
5. Byrne, H. M. and Chaplain, M. A. J. (1995) Growth of nonnecrotic tumors in the presence and absence of inhibitors. *Math. Biosci.* **130,** 151–181.
6. Chaplain, M. A. J. and Stuart, A. M. (1993) A model mechanism for the chemotactic response of endothelial cells to tumor angiogenesis factor. *IMA J. Math. Appl. Med. Biol.* **10,** 149–168.
7. Sleeman, B. D. (1997) Mathematical modeling of tumor growth and angiogenesis. *Adv. Exp. Med. Biol.* **428,** 671–677.
8. Orme, M. E. and Chaplain, M. A. J. (1996) A mathematical model of the first steps of tumor-related angiogenesis: capillary sprout formation and secondary branching. *IMA J. Math. Appl. Med. Biol.* **13,** 73–98.
9. Perumpanani, A. J., Sherratt, J. A., Norbury, J., and Byrne, H. M. (1996) Biological inferences from a mathematical model for malignant invasion. *Invas. Metastas.* **16,** 209–221.

2. Word equations are a valuable first step before attempting a mathematical formulation. These should indicate the interdependencies between the various model variables, and it is also helpful at this stage to summarize the various data inputs available.
3. Mathematical equations are simply a rewrite of the word equations, but in terms of mathematical symbols. Typically all the terms are relatively standard, with mathematical novelty coming from the way in which they are combined.
4. Parameter identification is important for qualitative as well as quantitative modeling. The values of some parameters can be obtained directly from the scientific literature: Kinetic rate constants and chemical diffusion coefficients often fall into this category. Others will typically be determined using experimental data, often from in vitro experiments that led to the modeling study. It is often the case that some parameters cannot be estimated accurately, but that approximate or order of magnitude values are known. This is in no sense a barrier to effective use of the model, but indicates that it is important to perform a parameter sensitivity analysis, described in **item 6**.
5. Numerical simulation is a key step in the use of a mathematical model. A variety of approaches are possible, as discussed in **Subheading 2.** The best practice is to initially consider one particular set of biological parameters, and vary numerical parameters (such as spatial discretization) to confirm accuracy. The simulations can then be used as a "mathematical experiment," in which biological parameters can be varied to make quantitative predictions or test qualitative hypotheses, as appropriate. When model simulations run quickly on the computer, there is a temptation to rapidly generate a large volume of simulations that are difficult to handle; this can be effectively dealt with by identifying a particular series of "experiments" to do initially.
6. Parameter sensitivity analysis is not necessary in all cases but is often very instructive. The basic procedure is to vary individual parameters by a given amount (say 10%) and determine the resulting percentage change on a particular aspect of the model prediction. Ideally one should vary parameters in combinations as well as individually, but this is feasible only if the total number of parameters is relatively small. The purpose of this calculation is to determine whether the model predictions are particularly sensitive to one or two model inputs; if so, a more detailed study of these inputs may be appropriate.
7. Mathematical analysis is a powerful tool for the investigation of continuum models, as one can use analytical techniques developed previously in other areas of applied mathematics. Murray's book ***(14)*** provides a wide-ranging review of these methods. For discrete models, analytical tools are much less well developed, and one is usually limited to numerical investigation in these cases.

4. Notes

1. Numerical problems are usually manifested by simulation results that are clearly nonsensical (e.g., variables becoming extremely large, or negative), corresponding to the instability of the numerical scheme. Often this requires nothing more

10. Gatenby, R. A. and Gawlinski, E. T. (1996) A reaction-diffusion model of cancer invasion. *Cancer Res.* **55,** 4151–4156.
11. Redfern, D. (1996) *The Maple Handbook: Maple V* release 4. Springer-Verlag, New York.
12. Wolfram, S. (1993) *Mathematica: A System for Doing Mathematics by Computer.* Addison-Welsey, Reading, MA.
13. Leek, R. D., Lewis, C. E. Whitehouse, R., Greenall, M., Clarke, J., and Harris, A. L. (1996) Association of macrophage infiltration with angiogenesis and prognosis in invasive breast carcinoma. *Cancer Res.* **56,** 4625–4629.
14. Murray, J. D. (1989) *Mathematical Biology.* Springer-Verlag, Berlin.
15. Edelstein-Keshet, L. (1988) *Mathematical Models in Biology.* McGraw-Hill, New York.
16. Sherratt, J. A., Eagan, B. T., and Lewis, M. A. (1997) Oscillations and chaos behind predator-prey invasion: mathematical artifact or ecological reality? *Philos. Trans. R. Soc. B* **352,** 21–38.
17. Durrett, R. and Levin, S. A. (1994) Stochastic spatial models—a user's guide to ecological applications. *Philos. Trans. R. Soc. B* **343,** 329–350.
18. Weliky, M. and Oster, G. (1990) The mechanical basis of cell rearrangement. 1. Epithelial morphogenesis during fundulus epiboly. *Development* **109,** 373.
19. Ermentrout, G. B. and Edelstein-Keshet, L. (1993) Cellular automata approaches to biological modeling. *J. Theor. Biol.* **160,** 97–133.
20. Dallon, J. C. and Othmer, H. G. (1997) A discrete cell model with adaptive signalling for aggregation of *Dictyostelium discoideum. Philos. Trans. R. Soc. Lond. B* **352,** 391–417.
21. Morton, K. W. and Mayers, D. F. (1994) *Numerical Solution of Partial Differential Equations.* Cambridge University Press, Cambridge, UK.
22. Höfer, T. Maini, P. K., Sherratt, J. A., Chaplain, M. A. J., Chauvet, P., Metevier, D., et al. (1994) A resolution of the chemotactic wave paradox. *Appl. Math. Lett.* **7,** 1–5.