# Fighting Cancer with Math

# Module Overview for Instructors

Module description and purpose:

This module was created as a way to get students thinking in new quantitative ways using cutting edge breast cancer research as the context for discovery, letting students experience first hand the utility of math in science. The hope is that the clear import of the work, the "cool factor" of 3D imaging, and the personal story of the original investigator will engage students' attention and motivate their interest in analysis. Students follow along with a simplified but realistic approximation of the original research, and can make meaningful discoveries, even while in a computer lab or at home; ideally all our students could intern at biomedical labs to do independent research, but this kind of module can be a substitute when that isn't possible.

**Setting:** The setting comes from the work of [Dr. Kerri-Ann Norton](http://pages.jh.edu/~apopel/Kerri.html) at John Hopkins University. Kerri first used agent-based modeling to understand the major tissue types in Ductal Carcinoma In-Situ, and found an interesting result - the same tissue type (cribriform, looks like swiss-cheese) could be produced by two different mechanisms (or under two really different sets of conditions). She believed that because of the mechanisms at hand, the tissues might look similar in 2D but different in 3D. However, most tissue sampling looks only at thin slices.

Subsequently Kerri used 3D reconstruction of histological slices from patient biopsies to discover two tissue sub-types ("bubbles" and "tubes") that appear identical as 2D slices. Whether the two mechanisms of cell growth produce the two tissue sub-types is yet to be determined, but the links between modeling, empirical work, and clever image analysis are really compelling.

**Focal concepts**: This module can be used to teach algorithmic thinking, data visualization, frequency histograms, and image analysis concepts per se. More advanced courses could use this as a context for programming concepts and practice (i.e. by automating the point-and-click steps of image analysis with macros/scripts in ImageJ/ Matlab/ R).

## Alignment:

[**AP Biology**](https://secure-media.collegeboard.org/digitalServices/pdf/ap/ap-biology-course-and-exam-description.pdf)**:**

Learning objective 4.8: The student is able to evaluate scientific questions concerning organisms that exhibit complex properties due to the interaction of their constituent parts. [See SP 3.3; Essential knowledge 4.A.4]

[**Next Generation Science Standards**](http://www.nextgenscience.org/hsls-ivt-inheritance-variation-traits)**:**

|  |  |
| --- | --- |
|  |  |

Use a model to illustrate the role of cellular division (mitosis) and differentiation in producing and maintaining complex organisms. (HS-LS1-4)

Science and Engineering Practices: Asking questions and defining problems: Ask questions that arise from examining models or a theory to clarify relationships. (HS-LS3-1)

Science and Engineering Practices: Developing and using models: Apply concepts of statistics and probability (including determining function fits to data, slope, intercept, and correlation coefficient for linear fits) to scientific and engineering questions and problems, using digital tools when feasible. (HS-LS3-3)

Cross-cutting concepts: Systems and Systems Models: Models (e.g., physical, mathematical, computer models) can be used to simulate systems and interactions—including energy, matter, and information flows—within and between systems at different scales. (HS-LS1-4)

[**Common Core ELA Standards for Science and Technical Subjects**](http://www.corestandards.org/ELA-Literacy/RST/11-12/)**:**

[CCSS.ELA-LITERACY.RST.11-12.3](http://www.corestandards.org/ELA-Literacy/RST/11-12/3/)  
Follow precisely a complex multistep procedure when carrying out experiments, taking measurements, or performing technical tasks; analyze the specific results based on explanations in the text.

[CCSS.ELA-LITERACY.RST.11-12.7](http://www.corestandards.org/ELA-Literacy/RST/11-12/7/)  
Integrate and evaluate multiple sources of information presented in diverse formats and media (e.g., quantitative data, video, multimedia) in order to address a question or solve a problem.

[CCSS.ELA-LITERACY.RST.11-12.8](http://www.corestandards.org/ELA-Literacy/RST/11-12/8/)  
Evaluate the hypotheses, data, analysis, and conclusions in a science or technical text, verifying the data when possible and corroborating or challenging conclusions with other sources of information.

[CCSS.ELA-LITERACY.RST.11-12.9](http://www.corestandards.org/ELA-Literacy/RST/11-12/9/)  
Synthesize information from a range of sources (e.g., texts, experiments, simulations) into a coherent understanding of a process, phenomenon, or concept, resolving conflicting information when possible.

[**Common Core Mathematics Standards for Statistics and Probability**](http://www.corestandards.org/Math/Content/HSS/introduction/)**:**

[CCSS.MATH.CONTENT.HSS.ID.A.1](http://www.corestandards.org/Math/Content/HSS/ID/A/1/)  
Represent data with plots on the real number line (dot plots, histograms, and box plots).

[CCSS.MATH.CONTENT.HSN.Q.A.1](http://www.corestandards.org/Math/Content/HSN/Q/A/1/)  
Use units as a way to understand problems and to guide the solution of multi-step problems; choose and interpret units consistently in formulas; choose and interpret the scale and the origin in graphs and data displays.

[CCSS.MATH.CONTENT.HSN.Q.A.2](http://www.corestandards.org/Math/Content/HSN/Q/A/2/)  
Define appropriate quantities for the purpose of descriptive modeling.

[CCSS.MATH.CONTENT.HSA.CED.A.1](http://www.corestandards.org/Math/Content/HSA/CED/A/1/)  
Create equations and inequalities in one variable and use them to solve problems.

[CCSS.MATH.CONTENT.HSA.CED.A.2](http://www.corestandards.org/Math/Content/HSA/CED/A/2/)  
Create equations in two or more variables to represent relationships between quantities; graph equations on coordinate axes with labels and scales.

[CCSS.MATH.CONTENT.HSG.MG.A.1](http://www.corestandards.org/Math/Content/HSG/MG/A/1/)  
Use geometric shapes, their measures, and their properties to describe objects (e.g., modeling a tree trunk or a human torso as a cylinder).\*

**Instruction level**: Introductory or advanced undergraduate biology courses, high school, AP biology

**Prerequisite knowledge:** This module is designed to fit into a study of human biology, cell growth/physiology, histology, or the like. Alternatively, it could be used in a programming intensive course as a context for coding assignments. Previous experience with image analysis is helpful but not required for either the instructor or the students.

**Keywords**: data visualization, graphing, breast cancer, DCIS, histograms, centroid, sums of squares,

# Instructional approaches:

These materials are constructed to allow some instructors to just pick up the lab and go, and to allow others to construct their own customized lab from the pieces. Most instructors will find the complete lab to contain too many concepts, take too much time, and provide too much background. Just delete what you don't need.

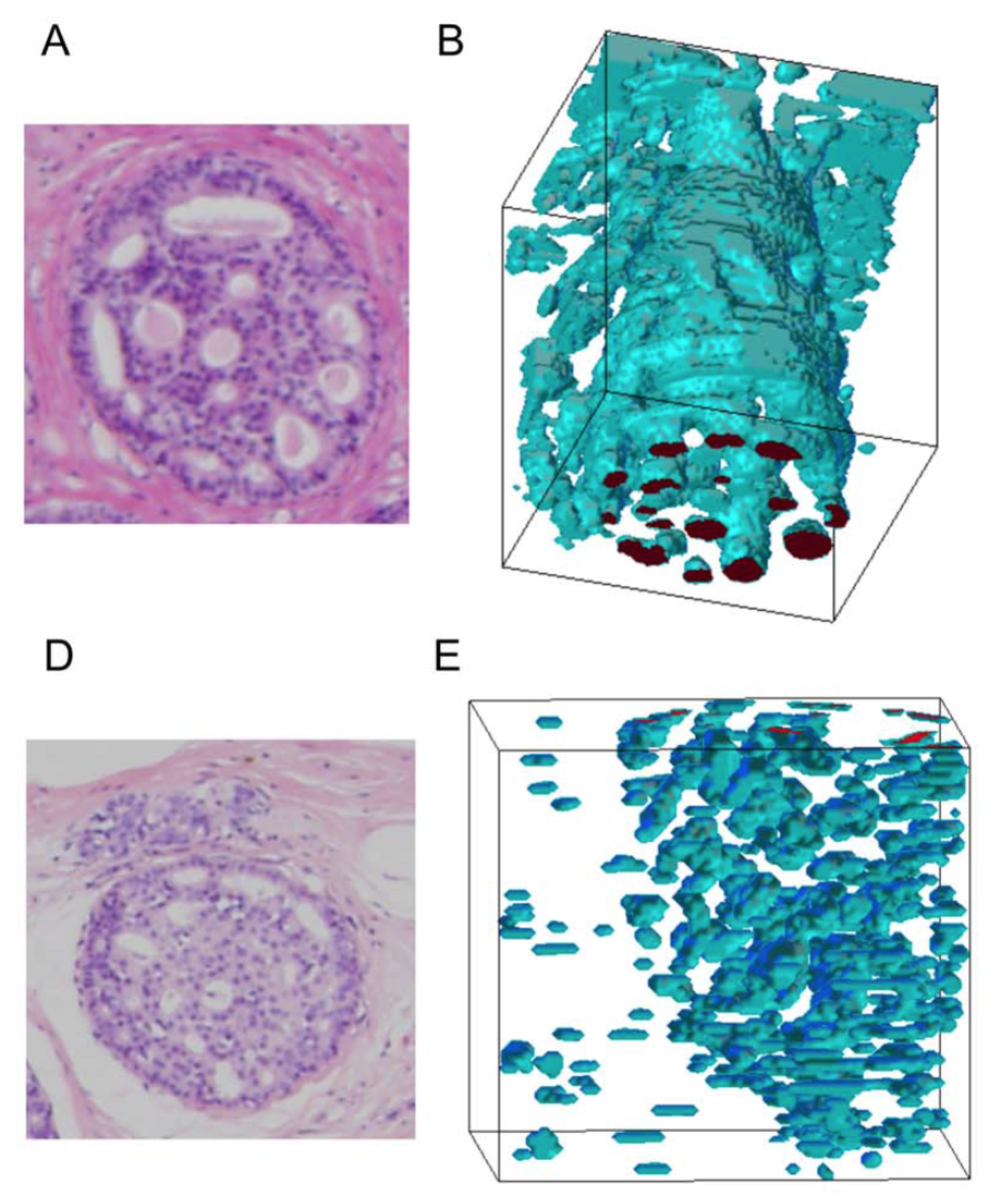
The image analysis/3D reconstruction part of the lab is *unfortunately* a lot of "cookbook" steps, following along the commands in ImageJ - the complexity of the steps makes it hard to be otherwise.

However, there is some potential for more open-inquiry for adventurous instructors - the order of the steps and some of the decisions that are made are not set in stone. Alignment can be done before segmentation (as is actually the case in Norton et al. 2012), and the degree to which any optional corrective smoothing is used is up in the air. An instructor could present the processes without a set order, and ask students to create the "best" 3D reconstruction they can... which forces them to think about what "best" means (visually pretty or smooth or easy to see? accurate? how do we know what is accurate when we can't see the tissue in 3D?). This is much more like the actual research process, where there is no "answer", and any method needs to be validated to be sure it isn't introducing artifacts. In terms of classroom mechanics, individual step (alignment, cropping, thresholding, etc.) instructions could be provided on separate sheets of paper, without an order and with less specificity.

**Discovery**: The real discovery part of this activity is finding two different 3D tissue subtypes among cribriform tissue... "bubbles" and "tubes", which are only revealed at the end of the process. Both look like swiss cheese in 2D, but the latter has much larger/longer continuous luminal spaces in the vertical direction. Below is an excerpt from Fig 6 in Norton et al. 2012, showing very similar looking 2D slices, but very different looking 3D tissue types.

The students are "prepared" for the discovery by the narrative and several animated simulations of an agent-based model that Kerri made in 2010. The animations show different kinds of tissue development.

**Ideally, student groups will perform 3D reconstructions, and then notice, identify, and explain these two tissue types**. This is the big pay-off for the lab, and value that was hopefully worth the cookbook, "click here" approach. However, students may not be used to having anything to discover... they may be much more used to finding some result, printing it out and turning it in, because "whatever I found is what the teacher knew was going to happen anyways." And, probably only very keen students will remember that this whole process started with Kerri noticing two different processes were leading to seemingly similar cribriform tissues (in 2D). The potential connection between two different growth processes and two different 3D tissue subtypes (bubbles and tubes) may need to be drawn out of students from prompts in post-lab assignments or in discussions.



**Part of Fig 6 from Norton et al. 2010b**

Moreover, the connection between earlier modeling and this empirical work also might slip past students, without prompts or discussion. That the modeling predicted two different processes, and then the clever 3d reconstructions found different subtypes is really a remarkable bit of science. We talk a lot about how the different major approaches to science (experiment, observation, math/modeling) work together, but this is an awesome example. Clever ways of having students "experience" this, rather than just be told about it, would be a welcome addition to this lab.

# Quantitative concepts:

**Basic**

* Basic graphing
  + Frequency histograms (distribution of cell frequencies across space, distribution of color values in images)
* Centroids
* Correlation
* Sum of squared deviations
* Programming in pseudo-code

**Advanced**

* Cross-correlations with normalization
* Programming
  + automate image analysis steps in macros/scripts
* Measuring 2D dimensions from 3D reconstructions

# Potential learning objectives:

**Basic**

* Students will be able to use sums of squares to describe differences.
* Students will use image analysis software to generate data from an image set.
* Students will use image analysis software to generate a 3D data visualization from an image set.
* Students will be able to construct and interpret frequency histograms.
* Students will gain appreciation for the interplay between modeling and empirical work.
* Students will gain appreciation for the utility of mathematics and computation for biology and medicine.
* Students will write a discussion section in typical scientific literature format.

**Advanced**

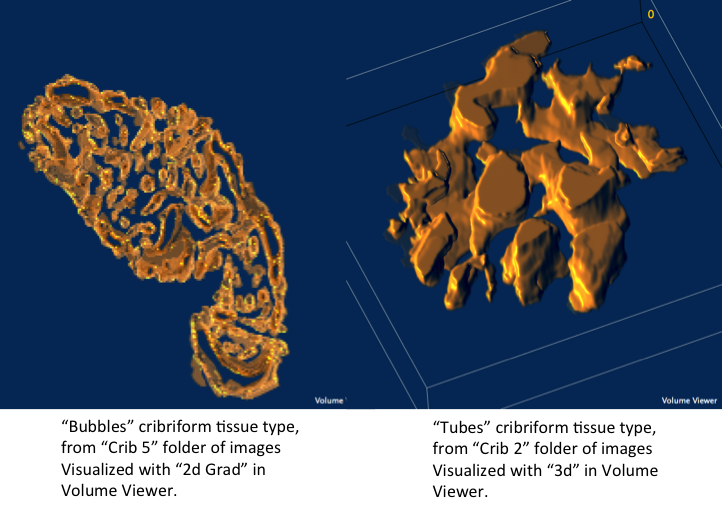
* Students will apply programming concepts to automate a series of procedures.
* Students will quantify aspects of 3D shapes using image analysis.

# Details on the Main Lab Activity:

**Introduction/Pre-lab:** The first three pages are background on the subject matter and set the scene for investigation. These pages would be ideal for a pre-lab reading assignment. Having students complete the first question, where they are asked to consider what steps will be necessary to perform 3D reconstruction from a stack of images, would also be good as a pre-lab assignment, and could form the basis of a brief discussion at the beginning of lab. The point of this question is for the students to have some framework or scaffolding on which to hang the steps they will be asked to perform. Without a larger view of the process, they may quickly get lost in the "point here" and "click here" nature of the instructions.

**Lab Activity**: Most of the steps in ImageJ are pretty straightforward, and most problems will just stem from incomplete/inaccurate reading of the instructions.

However, the selection of options and viewing angles during the 3D reconstruction can be a bit tricky... some students may get off track and have trouble. Moreover, some image stacks "look better" as 3d reconstructions using the "2D Grad" option on the Volume Viewer, and some work better using the "3D" option. See below for an example... unfortunately, it seems like "bubbles" and "tubes" are best visualized under different conditions. (Remember that the "solid" depicted here is defined as the luminal space, so this is kind of a "negative" view of the duct).



Moreover, there is no "right" angle to view from, and no biologically defined color/shading/lighting that is "correct"... there is a bit of art to it, and having students think about what this means scientifically could be really interesting. We often forget about the degree of artistic rendering there is in any reconstruction, whether it is a fully-fleshed Neanderthal at a museum, a skeleton reconstructed from bone fragments, or tissue like those studied here. (This is also a potential source of frustration for students who tend to like very cut and dry instructions and outcomes. The 3d reconstructions can look VERY different depending on the parameters chosen).

If the reconstructions are a bit sloppy, distinguishing the two types may be challenging. The reconstructions pictured several pages above from Kerri's original work are >100 slices deep, which definitely helps make the vertical differences more apparent. Here, the stacks are not nearly as deep, to make the quantity of work manageable for students.

Besides the 3D reconstruction, just moving through the 2D slices back and forth gives a pretty good sense for 3D structure. A stack of .gif files can also be saved as an animated gif, which rotates through the slices, showing a sort of "time for height" substitution view.

When it comes time to evaluate student work, or decide if their decisions about tissue sub-types are accurate, remember that any tissue sample can be a mix of types. You can have a completely normal duct right next to a bubble sub-type cribriform duct, for example.

**Dividing up the work:** If the fully automated procedure is followed, students could likely quickly work through all six "patients". If a more manual approach is used, there are too many stacks, and too many ducts per image, to have each student or student group do everything. How many samples to do (either in terms of stacks (i.e. tissue samples) or ducts/stack) is therefore a function of class size, available time, and instructor preference. Having a class work together on two reconstructions, with one being a bubble and one being a tube sub-type, would presumably be the minimum.

If time/effort permits, quantifying the number of different tissue types and subtypes in each sample could produce some new, interesting results. In other words, does one patient sample have 100% bubbles, while another has 50% tubes and 50% bubbles? It would take a lot of work both by the instructor and students, but this could even be cast in a case-study framework... two different patients with the students as pathologists making diagnoses.

# Credits

Original research setting: Kerri-Ann Norton (and colleagues from 2010/2012 papers)

Help adapting research to module: Kerri-Ann Norton

Module content and writing: Jeremy Wojdak

Review and formative suggestions: