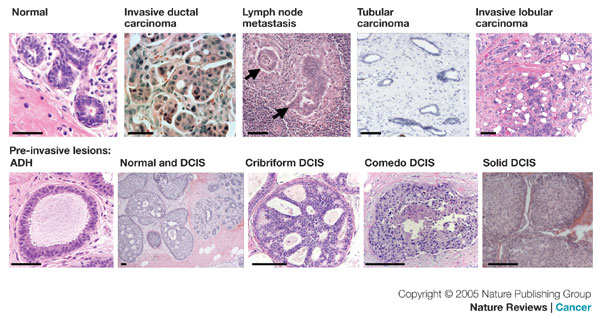
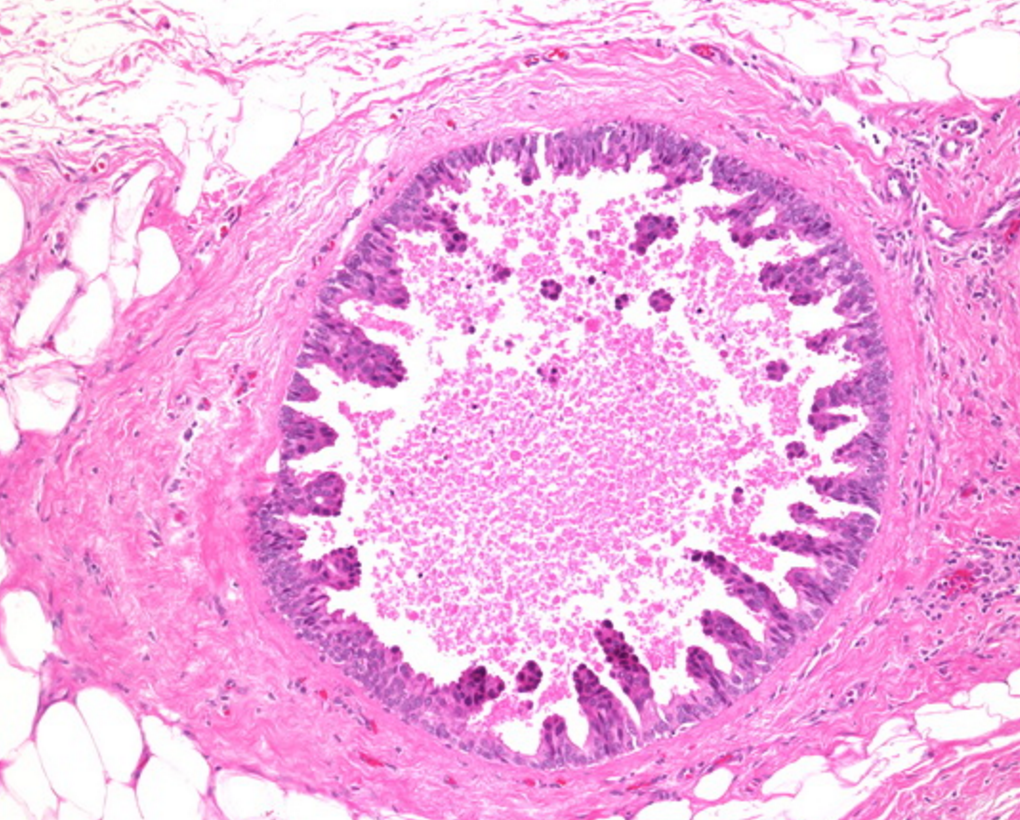
# Fighting breast cancer with mathematics

Breast cancer affects millions of families around the world, as one of the leading causes of cancer-related deaths for women. Breast cancer isn't a single disease, but really a variety of different kinds of abnormal cellular growth (see picture below, from [Debnath and Brugge 2005](http://www.nature.com/nrc/journal/v5/n9/full/nrc1695.html)). Different cancer types may have different underlying causes and thus have different prognoses and treatments.



One important type of abnormal cell growth is called Ductal Carcinoma *In Situ* (DCIS), accounting for about 13% of newly diagnosed breast cancer cases (Quinn and Ostrowski 1997). DCIS gets its name because of abnormal epithelial cell growth in ducts in the breast tissue, with the abnormal cells remaining inside the duct (hence *in situ*). Many untreated DCIS cases will progress to invasive cancer.

DCIS is diagnosed by biopsy and [histological](https://en.wikipedia.org/wiki/Histology) examination. Tissue is removed from a patient, fixed chemically, embedded in a block of paraffin wax to stabilize it, sliced very thin, stained with chemicals that differentiate cell types, and mounted to a microscope slide. These virtually two dimensional slices (only a few microns thick) of tissue are then examined. In DCIS, there are four main tissue types: micropapillary, cribriform, solid, and comedo ([Silverstein 2000](http://www.annualreviews.org/doi/full/10.1146/annurev.med.51.1.17), [Jaffer and Bleiweiss 2002](http://onlinelibrary.wiley.com/doi/10.1002/jemt.10180/abstract;jsessionid=DFC5C4CA402EA3733DD65E172E1D0985.f03t01); first pictured to right, latter three pictured above). Micropapillary has fern or branch-like extensions from the duct wall into the interior. Cribriform is "swiss-cheese" or sieve-like. Solid, is well, you guessed it, solid. And comedo is solid but with a necrotic (dead) core. These four tissue types are used to determine patient prognosis and potential for the cancer to invade other tissues. However, little was known about these four tissue types, how they formed, or whether they represented a progression of tissue development. Did they come from a single process or were different mechanisms required to generate the different tissue types? Obviously understanding how the tissues form would be a big step towards diagnosing and treating the disease.



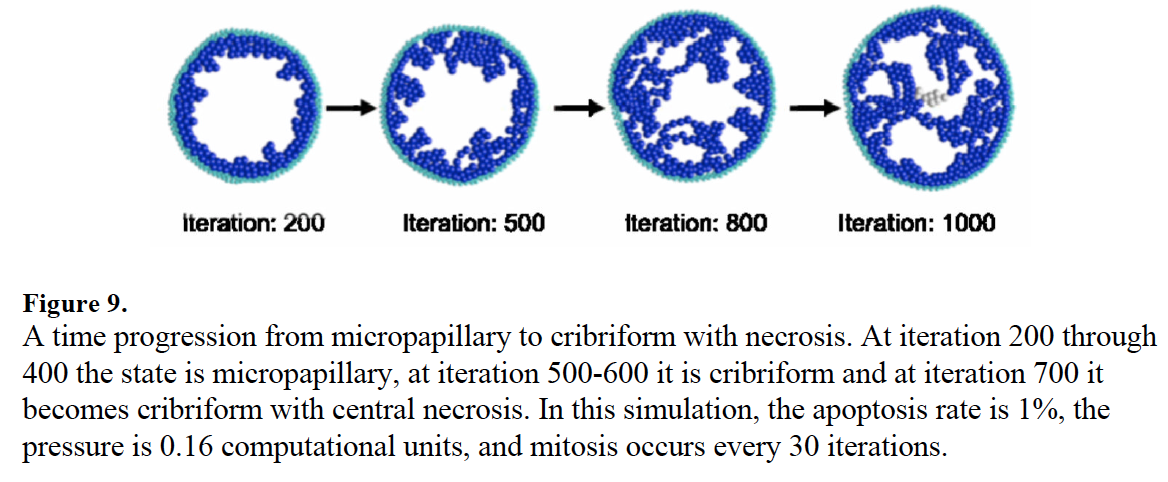
**Micropapillary**



Enter Dr. Kerri-Ann Norton. Dr. Norton, or Kerri as we will call her, was working at Rutgers University and turned her attention to DCIS. Kerri's expertise lies in biomedical engineering, computation, and mathematics - topics that you might think are only tangentially related to processes of cellular growth and medicine. However, as you'll see, math can be a potent tool in the fight against cancer.

In order to better understand the four main tissue types in DCIS, Kerri and colleagues created an [agent-based model](https://en.wikipedia.org/wiki/Agent-based_model) - a kind of computer simulation that would mimic the important aspects of cell growth and death. Agent-based models are used across science, from understanding bird flocks and flu epidemics to predicting traffic patterns - any time when the behavior of individuals controls the properties that emerge at higher levels of organization.

Kerri thought that by setting up simple rules for cellular processes like mitosis, adhesion, and death, properties at the level of the tissue (i.e. many cells) might emerge, and we might therefore understand more about how tissues form. In other words, could the same set of rules for cell "behavior" produce all the tissue types in DCIS? Would one type progress into the next, like a chronology? Or, would different processes at the individual cell level be needed to reproduce the patterns we see in real patients?

As you might imagine, both the methods and the results were somewhat complex and nuanced. You can read the original research paper ([Norton et al. 2010](http://www.sciencedirect.com/science/article/pii/S0022519309005669)). But just to whet your appetite, Kerri saw that different values for parameters like intraductal pressure, mitosis rate, and apoptosis (i.e. programmed cell death) rate could produce all the major tissue types. Moreover, when running simulations through time, one tissue type could transition to another. Below is a figure from that paper:

Here are several animations (they may take a minute to load) of simulated duct tissue growing under different parameter values. The light blue are myoepithelial cells that constrain the growth of the dark blue epithelial cells, and gray are dead cells.

[Low intraductal pressure (0.02), high apoptosis (0.02)](http://www.radford.edu/jmwojdak/AIMS_Cancer/animation1.gif)

[High intraductal pressure (0.14), low apoptosis (0.01)](http://www.radford.edu/jmwojdak/AIMS_Cancer/animation3.gif)

[Very high intraductal pressure (0.18), low apoptosis (0.01)](http://www.radford.edu/jmwojdak/AIMS_Cancer/animation4.gif)

What tissue types result at the end of each of these simulations? What is different/similar across simulations?

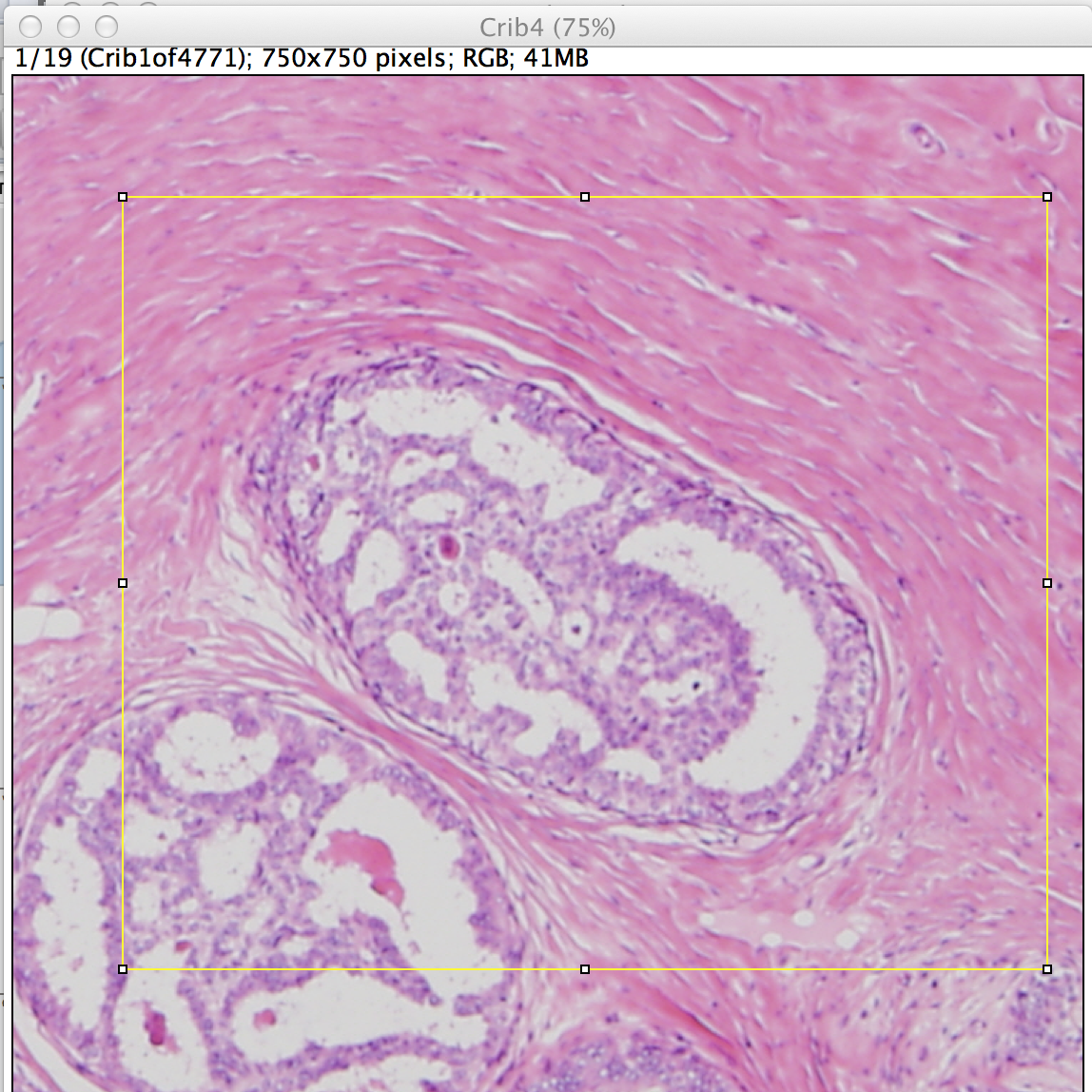
**One really interesting result struck Kerri as worth following up on - the cribriform tissue type was produced under two very different scenarios**. Cribriform tissue developed as a stable form with rapidly dividing cancer cells with high rates of apoptosis (programmed cell death), but also as a transient form with more slowly-growing micropapillae forming closed loops then opening up again. So it seemed as though two different processes were causing what might appear to be the same state, at least in thin slices of tissue, taken at a single point in time.

**Our investigation picks up with Kerri's at this point, as we follow along, recreating the steps she took**. Using different scientific approaches like experimentation, observation, and mathematical theory and modeling together is often very powerful, and can lead to insights not available to any single approach. From her modeling, Kerri had a hunch that reconstructing the cribriform tissues in three dimensions might provide new insights into the tissue formation. Kerri and colleagues decided to turn from the digital world of computer simulation to the empirical world of actual cancer tissue. They obtained images of serially sectioned (think of slicing a salami, keeping all the slices in order) tissue from DCIS patients.

Now, take a look at [**this stack**](http://www.radford.edu/jmwojdak/AIMS_Cancer/Crib2/) of images from serially sectioned tissue. What steps do you think we will need to take to reconstruct a single duct in three dimensions? Think very carefully, and work step by step. You don't need to know HOW to do each step (yet!), but try to think through what steps will be necessary. Detail the steps here:

# Image analysis

Let's see how well you did anticipating the steps required to reconstruct biopsied breast duct tissue samples in three dimensions.

1. **Get the images.** Go [**here**](http://www.radford.edu/~jmwojdak/AIMS_Cancer/)and select a folder that represents one set of slices of a single tissue sample. It is likely that your class will divide up the folders amongst groups, so be sure to work on the right one. "*File*", "*Import*", "*Image sequence*", choose folder of images. Creates a 'stack' of images that are now a single file, with many 'slices' that you can move through (notice scroll bar). Save your new file with a meaningful name, as a "*Tiff*" file.

2. **Select a duct.** Some images will show more than one duct, so you may need to select a particular duct to focus on. Then, to reduce file size and make things more manageable, you should crop the image around the duct you selected. Don't crop too close - leave a good amount of room around it - from slice to slice the duct may not be in exactly the same spot, and you don't want your crop to cut off part of the duct in some slices. Select the rectangle icon, draw a box, move the slider through the slices to check that the crop works for all slices, then select "*Image*", "*Crop*". **(Side note: throughout this process you may want to "*Save as*" after every major step is finished, using a new file name ("crib1.2.tif", "crib1.3.tif" and so on), so you can go back a step if you make some terrible mistake. Trust me).**

3a. **Isolate** (**'segment') the cribra.** Now, you need to isolate or 'segment' the cribra or holes in the duct. We will do this first manually. Choose the polygon tool (3rd from left on toolbar), click on the edge of a cribrum, then continue clicking along the edge, defining the boundary between cells and hole. You want to be accurate, but don't go overboard outlining every detail... this will only complicate the 3D reconstruction and you may be adding more noise than information. In particular, look out for extracellular debris (often stained pink rather than purple) that is just stuck in the hole... ignore that. Once you define the circumference, hit "*Ctrl\_x*" or "*delete*" to cut away that part of the image (should turn white). Use "*Ctrl\_+*" to zoom in and "*Ctrl\_-*" to zoom out, as needed. Proceed until several of the cribra in the duct are segmented.

3b. **Automatic segmentation**: To do this for all cribra in all the images can take some time... ok, I'll be honest... a LOT of time. We are going to be clever and ask the computer to help us out. What were we doing with our eyes? We were basically looking for high brightness/low color saturation areas that were reasonably continuous and contiguous... patches of "light". The computer can do this too, as long as we tell it what to look for. **GO BACK TO THE ORIGINAL IMAGE STACK, WITHOUT YOUR MANUAL SEGMENTED CRIBRA.** If necessary, re-import and save.

1. **Equalizing brightness/contrast**: Because we are going to automate things, the images have to be reasonably similar from one to the next. Cycle through them, looking for any that seem much brighter, darker, or more pale/intense than the rest. If you find any, adjust their brightness and contrast to make them more similar to the rest of the images. Use "*Image*", "*Adjust*", "*Brightness/Contrast*", then the sliders to manually change any unusual slices to match the rest. Alternatively, just hit "*Auto*" for each slice.
2. **Thresholding to outline cribra**: Now, we want to segment the cribra. Choose "*Image*", "*Adjust*", "*Color Threshold*". Then choose the "*Threshold color*" as "*B&W*", check "*Dark background*", move the sliders for saturation to choose "*0*" and "*50*", top and bottom. All the "*Pass*" buttons should be selected.

If you click back and forth on "*original*" and "*filtered*", you can see what its doing...

you can adjust the max value of saturation from 50 higher or lower to optimize the outlining of the cribra, and to minimize the creation of black areas that aren't cribra.

Hit the "*Stack*" button. What the heck just happened? We told the computer to make everything without much color saturation black, and make all the pink/purple/darker areas white. Voila! Automatic segmentation. Notice, though, it isn't perfect... we have little bits of black here and there that don't represent cribra (plus cellular material outside of the ducts). We will take care of that in a minute. **Remember to "*save as*" periodically!**

1. **Cleaning up**: Now we need to take care of extraneous black dots that aren't cribra.
   1. (If at this point, or any other point, any of your images happen to be reversed compare to the others, just select that slice, go to "*Edit*", "*Invert*", and say "*No*" to the pop up. If they all get reversed.. say "yes" to invert them all.)
   2. Make the image binary (or pure black and white): "*Process*", "*Binary*", "*Make binary*".

* 1. Now to get rid of little extraneous bits of black. We will throw out anything below a certain size, and keep big patches. Choose "*Analyze*", "*Analyze particles*". Type in "*150-Infinity*" into the size box. Choose "*Show: Masks*", and say "*Yes*" to the pop-up to apply to all slices. You'll notice a new file is created... be sure to save this one.
  2. The 150 value for minimum area of black patches to keep is just a starting point. Verify this is a sensible value by comparing the outlined cribra with the original images... are all the holes you can see visually properly captured?

**5a. Manually registering/aligning the images**: The serial sectioning of patient tissue procedure isn't perfect - not every image is exactly aligned with the last. Imagine taking fatty (squishy) tissue and working with slices 4 microns thick, and using a microscope to take photomicrogaphs... tricky business.

We can *'translate'* or move left/right/up/down any slice, and we can *'rotate'* or move clockwise/counterclockwise any slice to improve alignment between subsequent slices in the stack. Here we will do this manually, but later you will learn about how it can be automated, and later still about how math is used in that automation. In image analysis parlance, this process of alignment is called **'registering'** the images.

Start with the first slice, and then move to the second. Go back and forth a few times. Notice which way the image tends to jump, overall. Look at several particular cribra and see which way they seem to move. This is the amount you will need to either translate or rotate the image.

Select the second slice. Go to "*Image*", "*Transform*", "*Translate*". Click the preview button on and off to see the effect of the movement. Change the numbers (positive, negative, magnitude) to move the second slice about the correct amount to have it line up with the first slice. Click "*OK*", then when a second box pops up, click "*No*". If you click yes or just hit return, it will move all the slices this much, which won't help you. You will almost certainly have to do this process several times before things line up.

Now, if you have things lining up pretty well, but there seems to be a rotational difference, you can do essentially the same thing with "*Image*", "*Transform*", "*Rotate*" to rotate the second slice to align with the first.

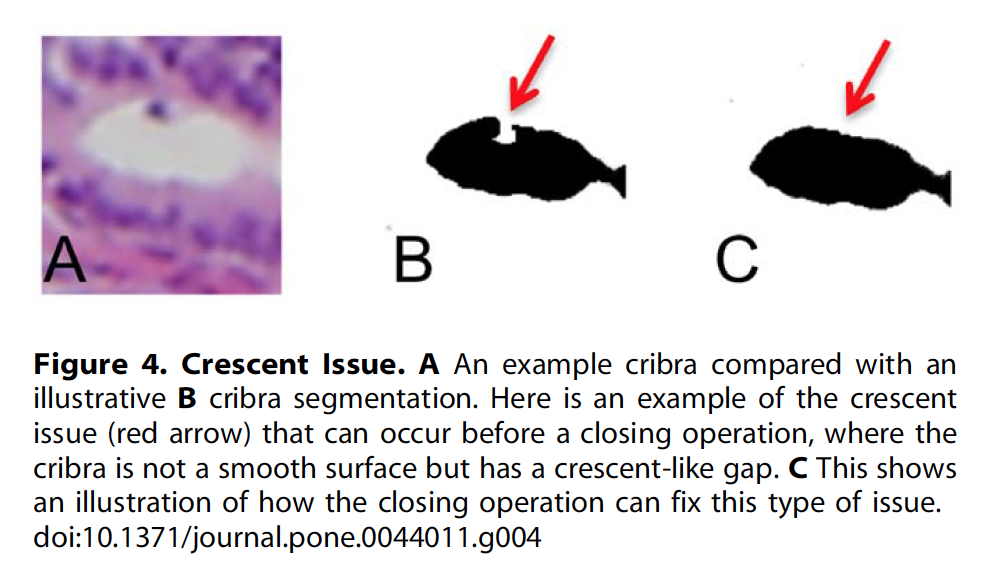
**5b. Automatically registering the images**: Once the first and second slices line up, you could move on to the third, and so forth. This would grow tiresome at some point, so we will again automate the process. We will also learn the underlying mathematical concepts the computer uses to accomplish this a little later.

There is automated registering method in ImageJ, as a plugin called "*StackReg*". If you choose "*Rigid body*", this will rotate and translate each image in the stack, without distorting the shapes to fit. It does a pretty darn good job, but go through each slice to make sure its as good as it can get. Once in a while one of the images may not be registered well... you can fix it manually. :-) If the colors get inverted, just go to "*Edit*", "*Invert*", and say "*Yes*" to do it to all slices.

{To use this plug-in, which itself depends on another plugin called "*TurboReg*",

download both [here](http://bigwww.epfl.ch/thevenaz/stackreg/), then put them (the .jar and .java files) in the "*Plugins*", "*Analyze*" folder in the file structure where ImageJ is installed on your computer. Restart ImageJ and it should be available via "*Plugins*", "*Analyze*", "*StackReg*". Your instructor may have done this ahead of time.}

1. **Optional smoothing, touch ups**. Really ragged edges, too much detail (especially if its not meaningful but just noise), holes in the holes, or other anomalies can all degrade the quality and interpretability of the 3D reconstruction. There are a few ways to clean things up. NOTE: Any of these steps could introduce artifacts into the reconstruction process, making the eventual image further from a "true" representation. There is a trade-off between truth and interpretability.
   1. My favorite approach is to dilate, fill holes, close gaps, then erode, in that order. There may be better approaches.
   2. We did this before, but just to make sure, lets make the image binary (or pure black and white): "*Process*", "*Binary*", "*Make binary*".
   3. Then make all the black patches a bit bigger... "*Process*", "*Binary*", "*Dilate*"...
   4. Fill holes in cribra - go to "*Process*", "*Binary*", "*Fill holes*".

e. Close tight gaps - if there is a strong indentation in an edge (see figure 4 from Norton et al. 2012 to the right), which happens often when there is cellular debris inside a cribrum, the outline may not be particularly accurate. Sometimes it can be closed or smoothed over by going to "*Process*", "*Binary*", "*Close*". Otherwise, you can choose the pencil tool and just draw it closed.

f. Then make all the black patches a bit smaller (back to their original size)... "*Process*", "*Binary*", "*Erode*"

Any of these steps can be repeated as necessary. The criterion that should be applied in making decisions is "Does this accurately capture the cribra I can see in the original images?"

8. **Secondary cropping**. Now that you've outlined the cribra, aligned the slices, and smoothed everything up, it is time to isolate just the duct of interest. Choose either the polygon tool or freehand tool (3rd or 4th from left on toolbar), and define a line that outlines the duct. The line should stay in white areas, and delineate inside the duct from outside. Once you do this, go to "*Edit*", "*Selections*", "*Make Inverse*". This just reverses what is selected from inside the selection to everything but the selection. Hit the "*delete*" key. Now you should just have the cribra from inside your focal duct, and nothing else. **You'll need to do this for each slice**. Have fun! You are starting to appreciate the value of automation, huh?

9. **Quality check**. Open up the original stack of color, unmanipulated images. Put this side by side with your finished black and white stack. Compare, noticing where your protocol successfully outlined cribra and where it failed. Remember, the trade-off for automation is often sub-optimal results... if we picked the single best outlines, the single best parameters at each step, our results would be better, but it would take forever. Are there any issues that are particularly bad? So bad that we should start over, or fix them?

8. **Oh yes. It is time. What time? 3D reconstruction time**. Choose "*Plugins*", "*3D*", "*Volume Viewer*". Starting in the upper right, choose "*Volume*", then "*Tricubic smoother*", set the "*Z-aspect*" to "*4*". The z-aspect defines how thick each slice is, relative to the x and y dimensions. Choosing high values exaggerates the height of the slices.

Then move to the bottom right and click the box for "*Light*". Then move to the upper right-ish and click "*3d fill*". Now go to the left side and click the cursor in all of the white holes visible in the depiction of the slices (top, middle and lower black and white panels). Notice you are adding to the reconstruction now - you are telling the computer that the white areas are the stuff you want to construct. If you get a pop-up, just hit "*ok*" and keep clicking. Move the sliders for "xy", "yz", and "xz", and keep clicking on white areas. Keep clicking. If you accidentally click on white, and the whole image goes crazy... just "*alt-click*" on white. Now adjust the "*Global alpha offset*" slider up to improve the image.

If you click on the image you can click and drag to change the angle of view. Once you have a good angle, good lighting, and a complete reconstruction, click "*Save view*" to take a snapshot... in a sense this is the "result" or "data" from your reconstruction.

**WARNING**: Depending on the size of your images, and the speed/memory of your computer, the 3D rendering can be slow. Be patient. There is a box called "*sampling*" - making this "*1*" will speed things up, making it >1 will give you better looking results.

Some image stacks are better viewed using the "*2D grad*" option in the upper right... there is some judgment here, try both, and decide on which one looks more like a real object.

Pause and reflect on what you've done. Think through the steps you took to get here and reconsider why you did each step. Why did we even start to reconstruct these tissues? Now, what are you seeing? What does the duct look like in three dimensions? Remember you are seeing the "holes" or the lumen in the duct... almost like a negative of the tissue. How would you describe the shapes of the holes, or more properly, the luminal spaces in the duct?

# Student Assignment

Congratulations! You've come a long way, following along with cutting edge cancer research, using quantitative reasoning and a bit of technology to understand biology. That is awesome. Now, let's tie it all up in a nice bow.

As a group or a class you've produced several 3D reconstructions, and now you have to determine what, if anything, can be learned from them. Re-read the introduction to the lab, and re-watch the model simulation videos. Now compare the 3D visualizations for different ducts... what do you see? What does it mean?

Scientific papers usually feature a discussion section, which briefly recapitulates the rationale or context for study, the main results, and interprets the results in light of the current knowledge in the field. Your assignment is to write a discussion of your 3D visualizations, briefly summarizing/describing them, what you think they mean, and how the work you just did fits into other knowledge about DCIS.

You will therefore need to do a little bit of searching, reading, and interpreting of the scientific literature. There are many sources for guidance for that task, including THESE and others your instructor may point you towards.

Your discussion need not be very long, perhaps only 1.5-2 pages, plus references cited.

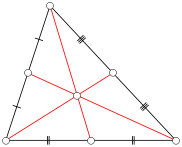
# Turning it up to 11: Advanced Assignment

You've learned a bit of image analysis. However, you can do many other things in ImageJ, including taking measurements from images. Now that you have 3D reconstructions, and can take 2D snapshots from them, you could conceivably take measurements of the objects in the reconstruction. For example, how long is a particular luminal space (cribrum)? What is the distribution of cribra lengths, or length to width ratios? Can you think of an interesting and informative way to quantify some aspect of the 3D reconstruction? Think about capturing some aspect of the heuristic view your mind takes in, and turning it into something numerical. Add a description of what you measured, why, and a graphical or tabular summary of that approach applied to several reconstructions.

# Side trip - automatic alignment

As you saw first-hand, the alignment of subsequent images in a stack is not trivial - it is hard to do well and its tiresome. Imagine doing that for ~100 images per sample, for many samples... ugh. Kerri used a computer algorithm and a bit of mathematics/statistics to handle the alignment automatically. Pause for a second and think about what you were doing when you were lining up images. What kind of comparisons were you making mentally?

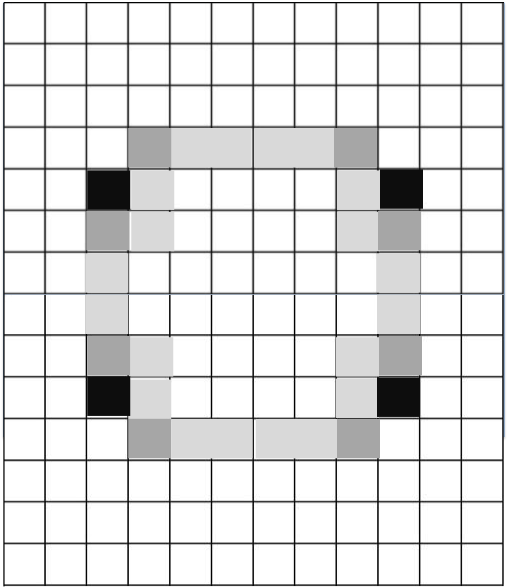
In what computer scientists call "pseudo-code", or code written in normal English, can you write a set of simple instructions for an automatic alignment procedure? What would the computer try to do and what would define "success"?

You can read about precisely how Kerri did it in [Norton et al. 2012](http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0044011), but here is a simplified version. She first selected a color or range of colors (purple-ish) that was indicative of duct tissue and used thresholding (like we did above) to isolate the main object of interest, the duct. Then she calculated a "centroid" or geometric center or average location of all points in the shape. To the right is an example of the centroid of a triangle.

Then, with a center identified, Kerri had the computer try different rotations around that center, of a second slice to best match up with a first slice. Well, we used our eyes to define "matching up", so what could a computer do? Kerri defined a cross-correlation... think of comparing each pixel between the two slices, and determining how similar they are. The rotation that maximizes the similarity between the values at each pixel location from one slice to the next is the best rotation to use.

Kerri then used a similar approach (cross-correlation) to determine the best translation, or movement up/down/left/right of a second slice to match up with a first slice. Voila! Automatic alignment, or "registering".

Let's follow Kerri's logic and calculate a simplified version of cross-correlation between image slices, and apply it to the translation stage of alignment.

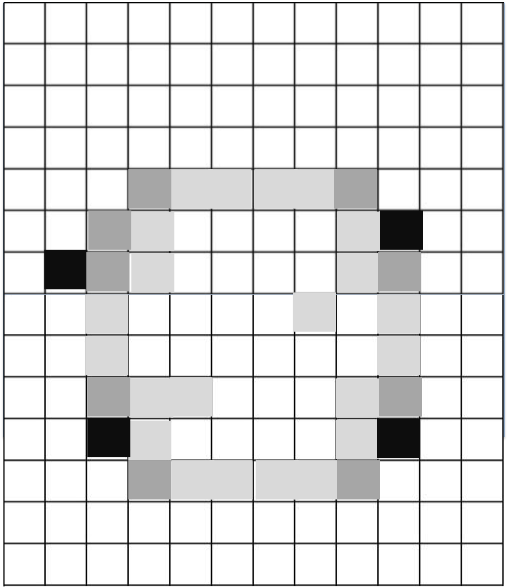
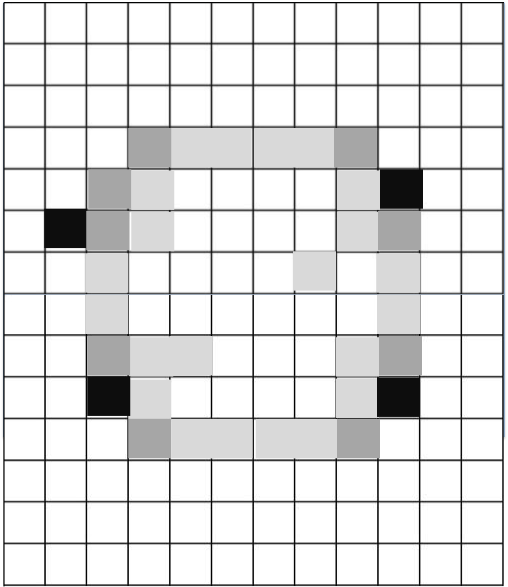
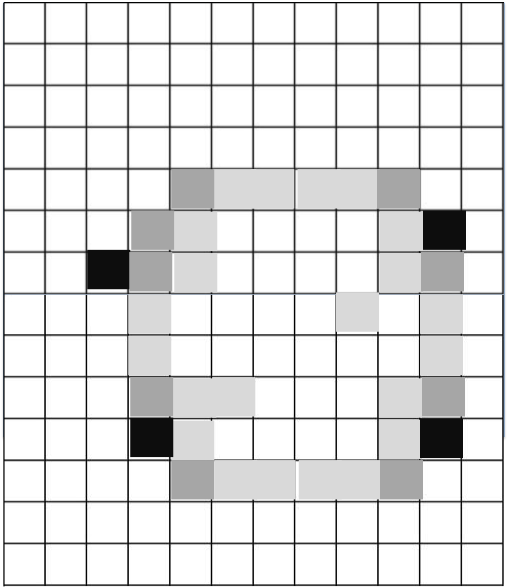
Here to the right is the "base" or first slice, to which we will try to align the second slice:

Now let's assign brightness values to the pixels. Let's consider the white spaces to be empty, give them a value of 0. Let's give the light gray a value of 1, medium gray a value of 2 and black a value of 3.

Now, let's define how well two slices align by the sum of squared differences between corresponding pixels. The more differences, the bigger that sum. Or, the better the alignment, the smaller the sum, if you'd like.

So for pixels that are empty in both slices, (0-0)2 = 0, and they don't contribute to our sum of differences. If the base image has a black pixel that is empty in the second slice, that is (3-0)2 = 9, and that contributes a lot to the summed differences. Why do we square the differences?

Now, take each of these three examples of potential new alignments for the second slice, and compare it with the base image - calculate the sum of squared differences in pixel values, across all pixels.



Which is the best alignment?

Now stop and consider what Kerri did again - instead of using brute force and her eyes to align 100's of images, she defined mathematically what she was doing mentally, and just told a computer to do it. This wasn't advanced calculus or something impossibly complex, but it was very clever. Math doesn't have to be hard to be really, really helpful. Can you write a general mathematical equation that represents the sum of squared differences among pixels along a grid? (Remember you can use subscripts like x, y to represent positions in a grid).

Now, if you were successful above, you have a way of quantitatively describing how good a particular alignment is. Can you write a pseudo-code algorithm to test out lots of possible alignments? What alignments should even be considered as "possible"? For example, should we even bother with considering an alignment that represents a 180 degree rotation of a second slice? or 181 degrees?

You'll notice that we registered our slices after getting rid of a lot of the image's information (after they were made binary). What effects do you think this would have on the registration process? Should we have registered the images first?

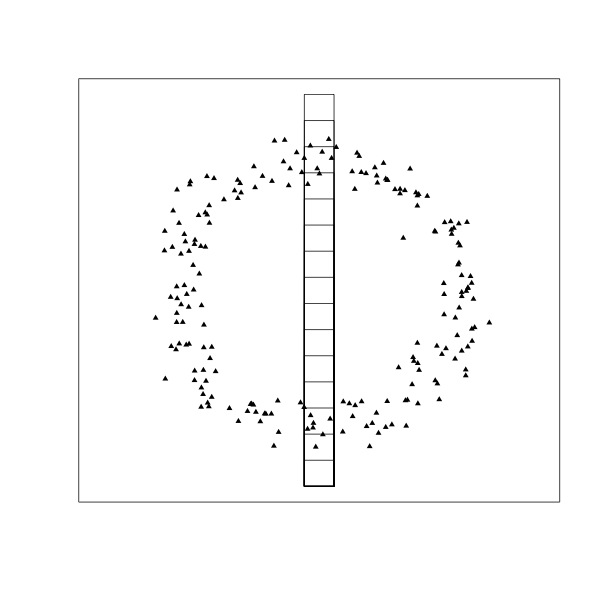
# Side trip - quantifying complex spatial arrangements, simply

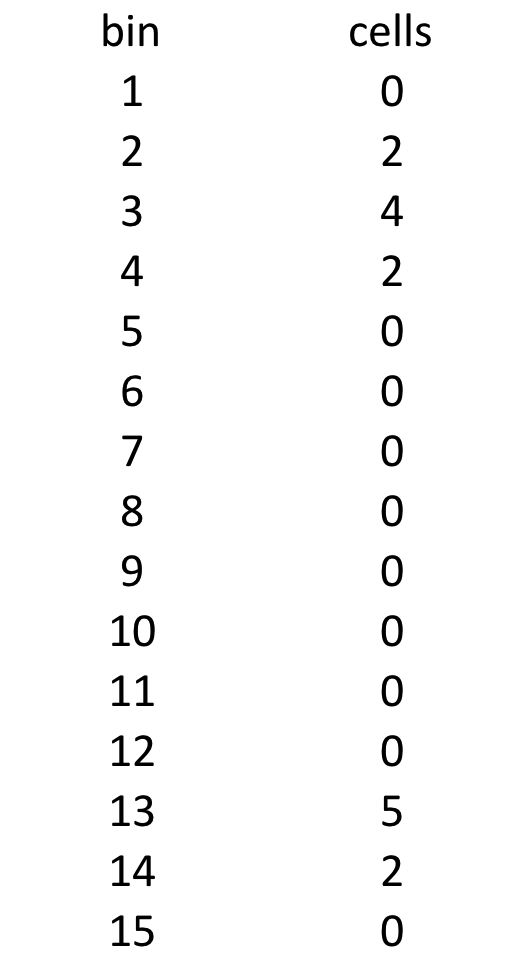
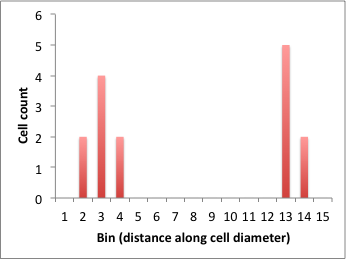
It is pretty easy to identify stereotypical, ideal images of tissue types like those in on the first page of the main handout (from [Debnath and Brugge 2005](http://www.nature.com/nrc/journal/v5/n9/full/nrc1695.html)). However, in many real tissue samples, the identification process is not always as easy, especially given that we expect transitions between tissue architecture types. Having reliable, *quantitative* ways of classifying and describing the morphologies is therefore invaluable.

Think about this analogy - when we see a pure-bred pitbull, nearly everyone can unambiguously identify it as such. But what about a half-breed? What about a mutt with a big square head? What are the traits that we are using subconsciously to recognize a dog as "a pitbull" or not? Having specific traits that we can quantify is clearly better than the more subjective, personal opinions of individuals, especially when identifying cancer tissue types (and therefore prognosis and treatment).

So, how can we quantify something as complex as the arrangement of dozens or hundreds of cells in a breast duct? Let's follow along with Kerri again.

Imagine we identified the center of a duct. Then, we placed little bins across the diameter of the more or less round duct, and counted the number of cells in each bin. Well, you might say that the number we get will depend a lot on where that diameter is placed... so we could do lots of them and take an average. Kerri did 100, using computers to speed up the process. We will just do one to get the idea. Now, how might we best display this data? Well, we could use a [frequency histogram](http://www.radford.edu/jmwojdak/histograms.mp4) that displays the frequency of occurrence in each of several (many) bins... this time, the bins represent distance along the diameter we overlaid. So, with the example here, I lay the bins across the diameter, count cells per bin,



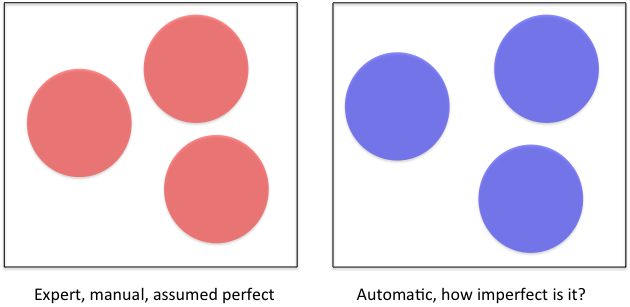


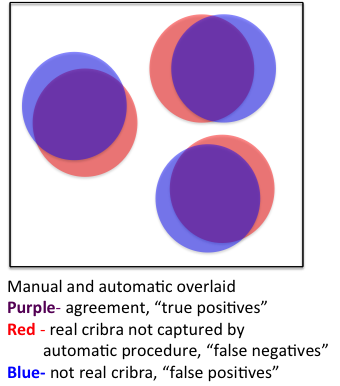
and to the right is a histogram of those cell counts... clearly "bimodal" or two humps. And as you can imagine, if we do more slices, they will all show that "two humped" pattern, because the middle of the duct is empty. Wow. We just came up with a nice, quantitative criterion with which to describe cellular arrangement in DCIS!

Now, follow the same logic, draw examples of the four main tissue types (e.g., micropapillary, cribriform, solid, and comedo) below. Then, draw an example diameter, count the cells in each of 15 bins, and make a frequency histogram, just like I did for the empty duct above. In essence, you are doing a thought experiment, and making quantitative predictions about the different tissue types. Let's see how well you do.

# Side trip - Evaluating our automated segmentation algorithm

One very important step in the procedure we used to reconstruct breast cancer tissues in 3D was the automated segmentation or outlining of the cribra. An obvious question is "How well did we do?" Kerri approached this question in her own work by comparing her algorithm to images manually segmented by expert cancer pathologists. If we assume the experts do a perfect job, than any differences between the automatic and the manual procedures can be assumed to be problems with the automatic one. Consider this visual example:





The relative rates or numbers of false positives, true positives, and false negatives are commonly used to assess the accuracy of medical tests, diagnoses, and procedures. For example, medical tests are often rated in terms of their sensitivity and specificity. Here we will use the same metrics as Kerri did, as described in Norton et al. 2012: "We compared the manually segmented images to the automatically segmented cribra using ‘‘precision’’ and ‘‘recall’’, two of many possible statistical metrics available. Precision and recall (Eq’s 2,3) are commonly used in engineering and segmentation research, and provide objective measures of how well an imaging algorithm agrees with ‘‘ground truth.’’ In general terms, precision is a measure of how many pixels chosen by the algorithm are outside of the lesion (false positives: FP) as compared with correctly identified lesion pixels (true positives: TP). By contrast, recall is a measure of how many lesion pixels are missed by the algorithm (false negatives: FN) as compared with true lesion pixels. In this way, precision and recall balance one other and often a high score in one implies a low score in the other.

Precision and recall are defined as follows:

**Applying the logic to our images.**

1. Choose one slice from the stack of images that you did a 3D reconstruction on. Proceed with the same steps you used before (except skip the automatic segmentation and do it manually for the entire image) until you get to a black and white image, with just black patches for the cribra. Save that as a file called "*Manual*". {This is like the red circle image in the above explanation}
2. Take the same image out of the stack from your finished automatic procedure, and save it. Open up your finished image stack, then choose "*Image*", "*Stack*", "*Make montage*", set "*Column*", "*Row*", and "*Scale factor*" to 1. Choose a single slice (that corresponds to the one you did manually), and put that slice number in the "*First slice*" and "*Last slice*" boxes. Save this image as "*Automatic*". {This is like the blue circle image in the above explanation}
3. Close down all your images. Open up the manually segmented image ("*Manual*"), and the automatically segmented image ("*Automatic*"). Both should be simple, binary images at this point (just black patches for cribra, with nothing else going on).
4. Go to "*Process*", "*Image Calculator*". We will use this submenu from here on out.
5. Set "*Image1*:" to the manual segmentation image ("*Manual*") and set "*Image2*:" to the automatic segmentation image ("*Automatic*"). Choose "*OR*" for the operation. Hit "*OK*".
   1. What you should see is black areas that show where the two images agreed - the true positives. {This is like the superimposed circle image in the above explanation, with purple areas highlighted}
   2. "*Edit*", "*Invert*" the image to get back to black patches showing the areas of interest. Save this image as "*TP*".
6. Set "*Image1*:" to the manual segmentation image ("*Manual*") and set "*Image2*:" to the automatic segmentation image ("*Automatic*"). Choose "*Difference*" for the operation. Hit "*OK*".
   1. What you should see is white areas that show where the two images differed. This is a combination of false positives (areas where the automatic algorithm assumed there were was cribra where in fact there was not), AND false negatives (areas where there was cribra but that the automatic algorithm failed to detect). We need to separate out these quantities and measure them. {This is like the superimposed circle image in the above explanation, with both pure red and pure blue areas highlighted}
7. "*Edit*", "*Invert*" the image to get back to black patches showing the areas of interest. Save this image as "*FP\_FN*".
8. Now, go back to the image calculator menu and set "*image1*:" to the manual segmented image, the operation to "*OR*", and "*image2*:" to the file you just created, "*FP\_FN*".
   1. You should see only some of the black pixels that were evident from the "*FP\_FN*" image. We've separated out the false negatives, or areas of cribra the automatic procedure missed. Save this as "*FN*". {This is like the superimposed circle image in the above explanation, with just pure red areas highlighted}
9. Now, "*FP\_FN*" was both false positives and false negatives. "*FN*" was just the false negatives. So, if we take the difference between those we should be left with the false positives.
   1. Set "*image1*:" to "*FP\_FN*", and set "*Image2*:" to "*FN*". Choose "*Difference*" for the operation. Hit "*OK*".
   2. You should now see just those pixels that the automatic algorithm thought were cribra, that were not in fact cribra... the false positives.
   3. Save this as "*FP*". {This is like the superimposed circle image in the above explanation, with both pure blue areas highlighted}
10. Now, to get a quantitative estimate of how well the automatic algorithm did, we can count the false positive and false negative pixels... the more there are, the worse the algorithm.
    1. "*Analyze*", "*Set measurements*", click on "*Area*" and "*Area fraction*"
    2. Open "*FP*", hit "*Ctrl\_m*" for measure. A results window should pop-up. Record or copy and paste the row of data into a spreadsheet or document. The "*Area*" describes how many total pixels there are in the image, and the *"%Area*" column describes what percent are white... the product of these two (divided by 100) is the number of white pixels, or the number of false positive pixels.
    3. Do the same for "*FN*", but you need to invert the image ("*Edit*", "*Invert*") before measurement. Record or copy and paste the row of data into a spreadsheet or document.
    4. Do the same for "*TP*".

**Problem Set:**

1. Now you have the data to go into the equations for precision and recall... calculate those metrics now. Report your results here, and discuss them in light of what Kerri reported about her own algorithm, Norton et al. 2012: "We determined that the median precision of the algorithm is 85.5%, with mean 74.9%... ; the median recall is 76.3%, with mean 71.2..."
2. Medical diagnostics are often rated in terms of **sensitivity** (what we learned above as recall), and **specificity**.

\* 100

**Specificity** is often described as the ability of a test to correctly identify patients without a condition. Notice the denominator is made up of all people who DON'T have the condition. If nearly everyone without the condition is diagnosed as being without the condition, the test is highly specific.

Consider sensitivity (or recall) again.

**Sensitivity** is often described as the ability of a test to correctly identify patients with a condition. Notice the denominator is made up of all people who DO have the condition. If nearly everyone with the condition is diagnosed as having the condition, the test is highly sensitive.

Ideally a test would be highly specific and highly sensitive.

1. Why do you think there is generally a trade-off between having a highly specific test and a highly sensitive one (i.e. many tests are really good at one, but less good at the other)?
2. Think as the patient. If you were told your test was POSITIVE, that it looked like you had the condition, would you hope the test was:
   1. highly specific
   2. highly sensitive
   3. not very specific
   4. not very sensitive
3. Think as the patient. If you were told your test was NEGATIVE, that it looked like you didn't have the condition, would you hope the test was:
   1. highly specific
   2. highly sensitive
   3. not very specific
   4. not very sensitive
4. Imagine we are studying a disease that is only found in 0.01% of the population, but is typically fatal within a month. We develop what we hope is a great test.... the test has 99% specificity and 98% sensitivity. Imagine we applied the test to 1,000,000 people.
   1. What number of true positives, false positives, true negatives, and false negatives should we expect?
   2. If you were diagnosed as positive by this test, you would:
      1. Sell everything, live like a king/queen, and wait for imminent death
      2. Shrug it off

Why would you make that choice (point to specific pieces of information from part a above)?

# References

[Debnath J, Brugge JS. 2005. Modeling glandular epithelial cancers in three-dimensional cultures. Nature Reviews Cancer 5(9):675-688.](http://www.nature.com/nrc/journal/v5/n9/full/nrc1695.html)

[Jaffer S, Bleiweiss IJ. 2002. Histological classification of ductal carcinoma in situ. Microscopy Research and Technique 59:92-101.](http://onlinelibrary.wiley.com/doi/10.1002/jemt.10180/abstract;jsessionid=DFC5C4CA402EA3733DD65E172E1D0985.f03t01)

[Norton K-A, Wininger M, Bhanot G, GanesanS, Barnard N, Shinbrot T. 2010. A 2D mechanistic model of breast ductal carcinoma in situ (DCIS) morphology and progression. Journal of Theoretical Biology 263(4):393-406.](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2839055/)

[Norton K-A, Namazi S, Barnard N, Fujibayashi M, Bhanot G, et al. 2012. Automated reconstruction algorithm for identification of 3D architectures of cribriform ductal carcinoma in situ. PLoS ONE 7(9): e44011. doi:10.1371/journal.pone.0044011](http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0044011)

[Silverstein MJ. 2000. Ductal carcinoma in situ of the breast. Annual Review of Medicine. 51:17–32.](http://www.annualreviews.org/doi/full/10.1146/annurev.med.51.1.17)