

PLACENTAL MATURITY ANALYSIS

SARAH GREFE, SGREFE@CSULB.EDU
RAYMOND AHN, RAYMONDJAHN@GMAIL.COM
MARIANA REYES, MARIANA.REYES@STUDENT.CSULB.EDU
JOSE RUIZ, JDRUIZ@LIVE.COM
DEPARTMENT OF MATHEMATICS AND STATISTICS
CALIFORNIA STATE UNIVERSITY, LONG BEACH

ABSTRACT. Placental maturity is a qualitative measure of villous evolution assigned by histologists, which reflects the uterine environment’s condition and fetal development stage [6, 3]. We quantified the geometric differences that constitute maturity, immaturity, and hypermaturity, by comparing the perimeter, area and circularity distributions of villi from microscopic slide images of placentas in each maturation category. A MATLAB program using Shai’s algorithm was created to remove maternal blood and segment images, which was performed with varying degrees of success on the histology image set. Another algorithm was produced to record area and perimeter measurements on the cleaned images, which we used to compare the performance of Shai’s algorithm against “perfectly-cleaned” benchmark images altered in Photoshop CS5. No significant linear relationships or correlations were found between medical variables and measurement variables, due to a small sample size (27) and relying on image quadrants to represent each sample.

1. INTRODUCTION

As the conduit between mother and baby, the placenta enjoys a complex status as a variable in the health of its infant. It utilizes the uterine environment, which is influenced by the mother’s health conditions, weight gain over the course of pregnancy, her nutrition and drug intake, and a number of other factors [3]. The placenta exchanges the substances diffused through the placental barrier to the fetus [3]. The placenta is thus a reflection of and influence on infant health, development, and gestation, and the degree to which it impacts fetal development is a challenge to gauge. Moreover, in order to support fetal growth, the placenta undergoes its own patterns of development which mirror the paths of change of the infant [3]. Consequently, conditions which alter or inhibit the placental growth patterns have an immediate impact on the fetus [3].

The mode of study which makes all of these influential factors on the placenta most apparent is histology. A microscopic slide of a placental cross-section can tell the trained histologist the health conditions of the mother, such as pre-eclampsia, diabetes mellitus, or abnormalities in the uterus; and about her external environment, such as if she lives at high elevation [3, 4]. A histological slide, most importantly, reflects developmental milestones in the fetus, through an aggregate of histologic observations united under a concept called Placental Maturity [3, 4, 9].

1.1. Defining Placental Maturity. Placental Maturity is a term histologists use to indicate a the degree of deviation from a normal range of visual features shared by placentas paired with healthy infants [3, 9]. The feature set for a healthy, term infant not only insists on the absence of disease (i.e. inflammation, edema, meconium stain, or hypoxia), the range of quantitative geometric and numerical properties of villi shape, blood vessels, syncytial knots, and fibrin deposits avoid extremes [3, 9].

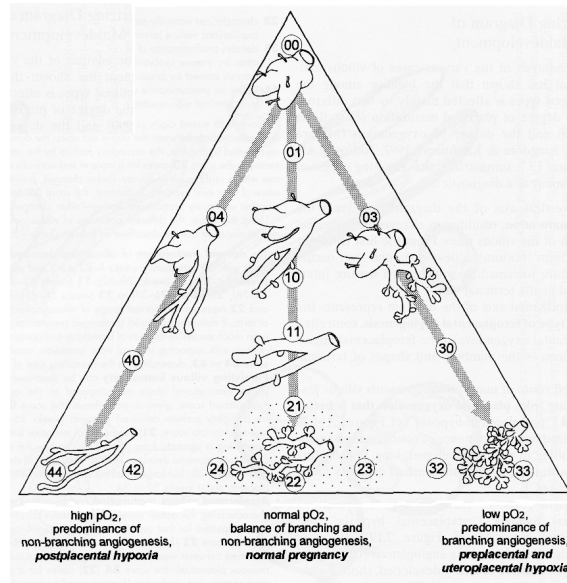


FIGURE 1. Diagram of villous developments and maldevelopment from [3]. The middle path represents normal development; the left path leads to immaturity, the right to hypermaturity. These paths represent respective extremes in the shapes of villi.

The placenta, as mentioned, follows its own growth patterns, and as the name implies, placental maturity for healthy infants greatly depends on chronology. Figure 1 succinctly shows the development of the placental villi, with paths to under or over development also displayed. Shown is the process of the villi beginning as a round ball, then branching (intermediate villi), then budding (terminal villi) [3].

Although ours is a simplified description of placental development, it focuses on what histologists expect to find in a cross-sectional slide of a healthy term infant's placenta: a non-extreme distribution of both larger, rounder intermediate villi and smaller irregularly-shaped terminal villi, with a reasonable number of blood vessels, syncytial knotting, and fibrin deposits [3, 4, 9].

Over the time of gestation, if one could ethically or physically observe weekly placental cross-sections, healthy growth would follow a progressive pattern as it matures [3, 9]. One round villi would eventually grow in size and split into many, smaller, irregular villi, until the image is congested with cross-sections of many small terminal villi and some intermediate villi [3]; simultaneously each week, more blood vessels would appear [3, 9]; the blue-staining nuclei of the epithelium-like tissue surrounding each villi (called syncytiotrophoblast) would begin as evenly spread across each surface, then in later weeks break up and gather into clusters (called syncytial knots) [3, 9]; concurrently, clots of blood from vessels would pool and coagulate into large red masses (called fibrin deposits), which increase in number with advancing gestation [3, 4].

1.1.1. *Asynchronous Maturity.* Akin to the fetus, the placenta can also undergo maldevelopments, which may be a reflection of a pathology in the uterine environment [3, 2, 7, 9]. In this case the placenta's failure to occupy optimum area on the uterine wall, whether through forming an irregular shape or insufficient size impacts the fetal development [7]. Conversely placental maldevelopments may simply mirror the fetal maldevelopments [3, 9]. In either case, an assessment of placental maturity which is asynchronous with respect to the actual gestational age is cause for concern on the part of the physician [3, 2, 9].

Immaturity. Immaturity may be defined as a placental maturity appearance which is not advanced enough for the gestational age of the infant [3, 9]. The appearance will manifest as

large irregular, perhaps long villi, with few small, round terminal villi present [3, 9]. Such an appearance at term may indicate, cause or correlate with meconium contamination, overly long gestation, high birth weight, infection, or retarded growth [3, 9, 4]. Blood vessels do not follow a single trend for immaturity; many pathologies either restrict or over-develop the proliferation of blood vessels despite how the villi develop [3, 4].

Hypermaturity. Hypermaturity is then the appearance of advanced gestational age without being in advanced gestation[3]. The quick overdevelopment of a placenta may trigger premature delivery since the placenta produces hormones which trigger the labor process [3]. A hypermature appearance may also be a symptom of, indicate, cause or correlate with premature birth, preeclampsia, hypoxia, suspected placental inefficiency [6], anemia, or gestation at high altitude [3].

In summation, placental maturity is a gestalt term used by histologists to describe a set of features for which a trained eye can instinctually recognize, as most humans can tell from a person’s face roughly what age they are. Qualitatively, most people could place a quantity (age) on a face and be within a reasonable number of years. We propose that placental maturity is a continuous measurement range of geometric properties. Histologists have only placed placental maturity into categorical classifications because there has so far been no definitive quantitative interpretation of the concept; thus the boundaries of the categories are also ill-defined. Indeed if a histologist were given a borderline case, some criteria should exist for them to decide how to classify the case. In principle, some quantitative thresholds should exist in order to validate maturity classification. The following paragraphs describe our geometric interpretation of histologists’ descriptions of placental maturity.

Villous Irregularity. We have mentioned the term “irregularity” in regards to the outline shape of villi. We have interpreted this as a deviation from roundness. Toward the end purpose of analysis and comparison between placentas, we have selected the shape factor *Circularity*, which measures how close to a perfect circle a shape is. Its values occur on a normalized scale where 1 represents a perfect circle, and numbers close to 0 represent eccentric or convoluted shapes [1]. A similar shape factor was used in reference [2], specifically $perimeter^2/area$. However due to the ability to directly measure circularity with Photoshop and the attractiveness of a normalized scale, we found circularity a more natural choice. To account for differences among images we later encountered, the independence from measurement dimensions aids to overcome potential errors. Mathematically [1]:

$$(1.1) \quad \textit{Circularity} = 4\pi \left(\frac{\textit{Area}}{\textit{Perimeter}^2} \right)$$

We expect the distribution range of circularity present in a healthy term placenta to be, on average, a value from 0.6-0.8, due to the mixture of intermediate and terminal villi, but predominantly terminal villi [3, 9].

Number of Blood Vessels. The number of blood vessels present indicates the density of blood vessels given a placental cross section [9, 4]. Immature placentas will have fewer, perhaps larger vessels, and hypermature placentas possess many small vessels crowding the villi [3, 9].

Number of Fibrin Deposits. The number of fibrin deposits increase with advancing gestation. We expect more deposits in more advanced maturity, though they occur with a large variance in healthy placentas [3].

Number of Syncytial Knots. The occurrence of syncytial knots also increase with advancing gestation [3, 9]. The knots also appear in varying degrees for healthy placentas, and a strong trend is not expected [3, 9].

2. RESEARCH METHODS

2.1. Data Set. Our sample set consisted of 27 unique subsets of mother, infant, and placenta. The histology images given for each placenta were initially scanned from one large cross-sectional

slice, at unknown depth, in the plane using the insertion of the umbilical cord as the normal vector. They were all colored using hematoxylin and eosin stain (H & E)[4]. However, due to logistical obstacles, we only gained access to images of smaller sections of the full-size images. This resulted in 64 images total, of which one placenta may have been represented by anywhere from 1 to 9 quadrant images. Our group also had no knowledge of how the images were selected for us; it is possible that a disproportionate amount of placentas with pathological traits, or traits “of interest” were given. We were also given no information on microscopic resolution, zoom or scale.

Also provided was a Microsoft Excel data sheet from a study performed by researchers at the University of North Carolina. It included an extensive list of variables and measurements recorded about the mother, infant, and placenta, and conditions surrounding the pregnancy (united under a single “Lab ID”).

2.2. Image Analysis to assess Placental Maturity. Our aim was to collect measurements of the relevant placental maturity variables, and analyze the UNC data to explore if placental maturity plays any role with medical variables. We then explored several avenues toward measuring images, through algorithm-operated and human-operated methods.

The following shows the process by which we cleaned, measured, and analyzed the images.

- (1) Clean-up the maternal blood and small villi of the image.
 - Image segmentation through MATLAB.
 - Image segmentation through Adobe Photoshop CS5.
- (2) Measure the perimeter and area of the image villi.
 - Measure MATLAB-cleaned image with MATLAB functions.
 - Measure Photoshop-Cleaned image with MATLAB functions.
 - Measure uncleaned image with Photoshop Selection tool method.
- (3) Count the number of blood vessels from the uncleaned images.

2.3. Measurement Variables. As mentioned above, we only measured the number of blood vessels per slide, villi areas and villi perimeters. There are two variables related to placental maturity which could not be analyzed through the programs we developed.

Blood vessels and fibrin deposits could not be counted automatically (or separately). First, since fibrin deposits, vessel tissues and red blood cells stain the same colors in the H & E dyeing process [4], the measurement methods could not distinguish between them, since the initial segmentation algorithms were based on identifying a “red” color. Second, even if an algorithm could identify and delete only fibrin and red blood cells external to villi, blood vessels in villi (and in general) consist of two types: arteries and veins [4]. Arteries must withstand higher pressure from the heart than veins, and the vessel tissues are resultantly thicker [4]. Vein walls are thinner, and in advanced placenta slides, may appear to be empty white space, perhaps filled with some red blood cells [3, 4]. The prospect of programming a computer to distinguish between all cases was too great for the time constraints and experience given.

A similar problem occurs when attempting to count syncytial knots. The distinction between an oblong trophoblastic border around a villi, a single syncytiotrophoblastic nucleus, and a true knot was too fine a distinction to implement in our algorithm, especially without scaling information that could be corroborated with available medical data.

The large variance mentioned in sections 1.1.1, and mentioned in [3], seemed to be too statistically weak to hypothesize that they would help correlate with any medical data. We thus found it satisfactory to use area, perimeter, circularity, and the number of blood vessels as the functional measurements for placental maturity.

2.4. Image Pre-Processing. Our goal in this step is to remove extraneous information from the histology image, but that is the final part of a multi-step problem. As seen in figure 2, there are maternal red blood cells present in the intervillous space that will interfere with future

measurement. Consequently, there are several things that must be done before the image is cleaned. First, we must be able to differentiate between the different colors as well as the slight variations in each color as well. Second, we must find a way for MATLAB to distinguish any red blood cells in the intervillous space from a blood vessel inside a villus. And finally, we need to be able to remove outside blood while keeping the inside vessels intact.

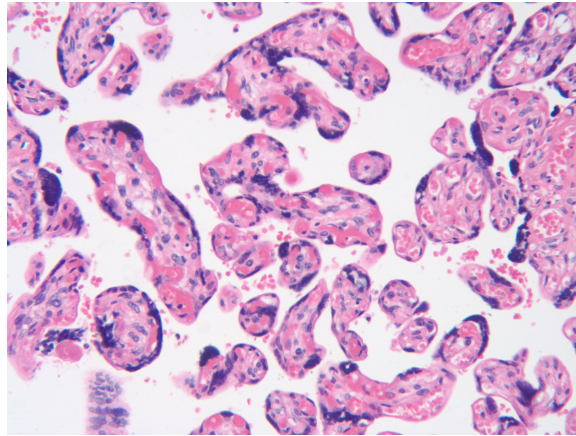


FIGURE 2. Original Histology Image 1479_1.

Because our results depend on measuring area of the villi as well as perimeter, any extra blood and cells outside of the villi will skew the results. So it is important that these elements on the image be removed before doing any calculations.

2.4.1. *Segmentation Through MATLAB.* Any image can be viewed as having a 3×1 vector for each pixel. This vector contains information for red, green, and blue color values. For example, a red pixel would have the vector $[1 \ 0 \ 0]$ while a blue pixel would have the vector $[0 \ 0 \ 1]$. Also any combination of these values from 0 to 1 would result in different colors being formed.

In figure 2, there are clearly four different colors: pink, red, purple, and white. Ideally we would separate the four colors and process each image separately, but we cannot just tell MATLAB to remove all the non-pink pixels from an image. If one were to carefully inspect the image, it is clear that for pink, and other colors, there are slight variations in the pixel value. So instead of just recognizing what is pink or red, we must have MATLAB recognize what is “pink-ish” and “red-ish.” That way all the colors that appear to be a certain color will be recognized as that specific color.

To deal with this problem, we considered two different methods of image segmentation: an automatic approach, and a non-automated approach. The automated algorithm we tried and had the most success with was based on Bagon Shai’s algorithm to quantify color images using mean-shift based quantizer [8]. The images in figure 3 were segmented using Shai’s algorithm three times.

Although the algorithm performed well in the far-left and middle pictures of figure 3, not all the colors are segmented correctly. Neither image contains the important red vessels. The final image is completely unusable for any calculations. If our goal was to simply isolate the villi, this algorithm would be useful, but since we want to calculate areas of each color, we need another way to segment the images.

The second method we tried, and eventually used, was inspired by a K-means segmentation with Mahalanobis distance metric [2]. Unlike the automated algorithm, this method relies on predetermined color markers. This way, instead of having varying results, all images will be transformed into a simpler image based on those color markers. For our images, we decided that an average pink value was about $[1 \ 0.7 \ 0.8]$, red was about $[1 \ 0.5 \ 0.6]$, and purple was about $[0.6 \ 0.4 \ 0.65]$. These numbers were obtained by manually selecting a typical color in one image

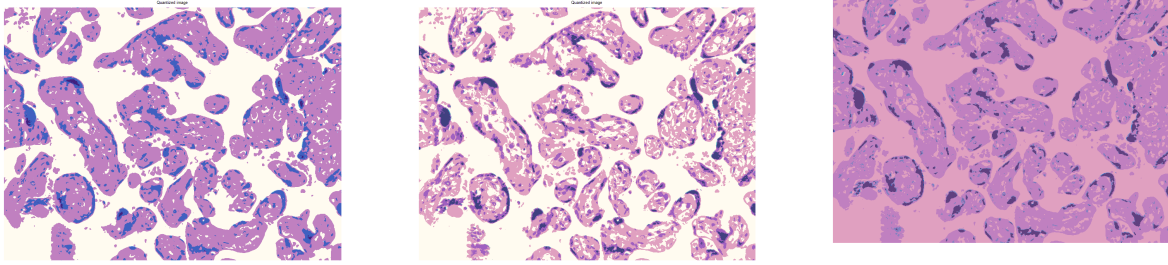


FIGURE 3. Automated algorithm output of slide 1479_1.

and averaging them throughout all our images. From here, we needed a way to determine if a pixel was pink or any other color. For this we simply used the Euclidean distance between a pixel value and the color markers and the closest color marker was used as the color value for that pixel. Using this approach, we were able to get a consistent result with a clear segmentation between our four colors.

The next thing we had to do was remove all the small isolated vessels that lie outside the villi. For this we simply used MATLAB's built-in morphological operations [5] to remove small areas under a certain threshold. The resulting image is as follows:

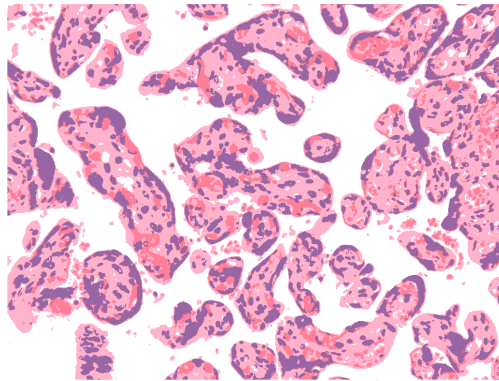


FIGURE 4. Predetermined Color Marker segmentation of slide 1479_1.

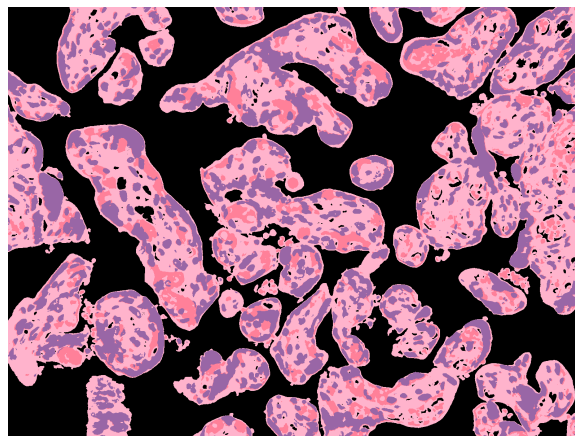


FIGURE 5. Cleaned image of slide 1479_1.

2.4.2. *Clean-up Through Photoshop CS5.* All of the original images we used for collecting our data were cleaned up using two different methods, a MATLAB algorithm we created and Adobe Photoshop CS5. “Cleaning” refers to removing maternal blood cells and fragments of villi in the intervillous space. In total, 52 images were “Photoshopped.” To perform the cleaning via Photoshop, two tools were used for every image, the Magic Wand Tool and the Paintbrush tool. For some images, a third tool called history brush was also used.

We used the Magic Wand tool to select the areas of the image we wanted to remove. Clicking on a pixel (which was considered the cursor pixel) in an image with the magic wand, all of the neighboring pixels which have equal value or deemed similar in color value are selected. These values can be specified through the options to determine the exact selection. The option we commonly used was Tolerance, which has values ranging from 0 to 255, based on the RGB color space. It determines how similar in color a neighboring pixel must be in order to be selected. Lower numbers imply a low tolerance and higher numbers imply a high tolerance. For example, if 0 is selected as the tolerance, all neighboring pixels will have to have exactly the same value to be selected; if 255 is chosen, all pixels will be selected. We chose the tolerance to be 20, 30 or 40 depending on the image. Then the selected pixels were removed by pressing the Delete keyboard command.

Colors that are similar which are not neighboring the cursor pixel will not be selected, leaving many fragments of cells which the tool could not identify. After using the magic wand tool as much as possible, we used the Paintbrush tool to replace the color of unwanted pixels with white pixels.

The history brush tool was used when the Magic Wand tool made accidental deletions. By opening the History palette, we could identify when the image was correct and paint the image back on. Figure 6 shows an example of an image that has undergone this process.

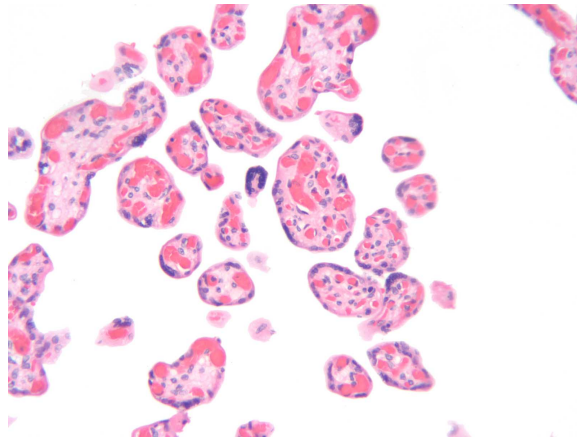


FIGURE 6. Photoshopped Image.

2.5. **Image Measurements.** The cleaned-up images obtained in the previous section are prepared for an algorithm to measure. The placental maturity variables we are interested in from the Photoshopped and MATLAB-segmented image sets are villi area and to a lesser degree, perimeter. As previously discussed, there were too many obstacles to effectively count the number of blood vessels.

The basis of the measurement algorithm we developed is counting, which presents problems later when attempting to use the data for analysis. This stems from the inclusion of villi that have edges extending beyond the edge of the image. The perimeter measurement included the straight border edges of villi; the area measurement of a single cut-off villus is similarly distorted. Any comparison we would potentially make using the circularity formula with these images

would hold no significance. For the Photoshop and MATLAB cleaned images, we designed the algorithm to simply count the total area of all villi, and the total perimeter of all borders.

In order to obtain reliable measurements that could be used for analysis of circularity, and additionally to measure the distribution of blood vessels, we utilized Photoshop CS5 and its measurement tools.

In summary, we collected two sets of total area and total perimeter measurements using a single MATLAB algorithm (one from the Photoshop cleaned images, one from the MATLAB cleaned images). For each non-edge villi from every original untouched image set, we collected area, perimeter, circularity, and number of blood vessel measurements.

2.5.1. MATLAB Measurement Program. In order to calculate total area and total perimeter from cleaned images, we implemented a MATLAB algorithm which measured the total area and perimeter of each image. The total area is the total number of pixels which are pink, red and blue. This also included all of the white pixels completely surrounded by pink, red or blue pixels, in order to account for empty veins inside a villus.

Total Area and Total Perimeter. We used one algorithm named Find_Areas (see Appendix section 2) to find the total area and perimeter for both sets of cleaned up images, Photoshopped images and images cropped the program Modeling_Project (see appendix section 1). The following outlines the details of the algorithm:

- (1) Imports an image from the specified file directory.
- (2) Converts all pixels of the image into pink, red, blue or white, depending on which color it is closest to in the RGB color space.
- (3) Segments the pink, red and blue pixels and calculates the area for each of them.
- (4) Combines the pink segmented image, red segmented image and blue segmented image.
- (5) Fills any area of pixels surrounded by pink, red or blue pixels.
- (6) Calculates the total area and perimeter.
- (7) Reads next image in the file directory and repeats until all images are processed.

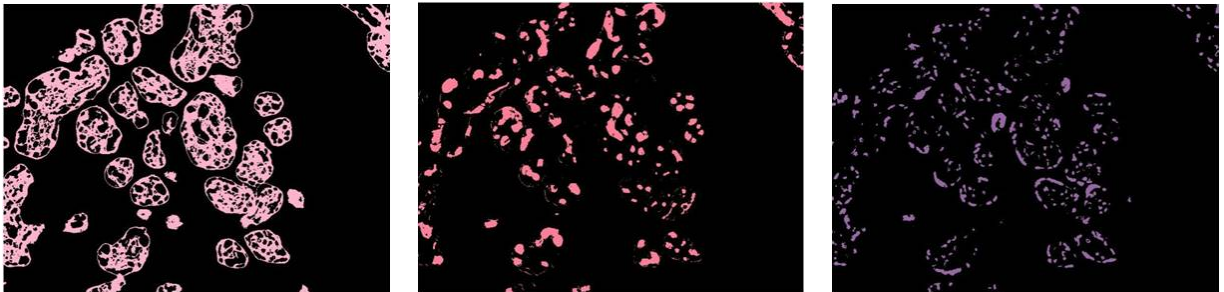


FIGURE 7. Images segmented by color values.

We used a simple loop to convert all the pixels of each image into pink, red, blue or white, depending on how each color value was defined. Through a process of trial and error we selected a set of color values that produced the best output for the greatest number of images. The loop considers the color value for each pixel, finds the user-specified color value the pixel most resembles, then assigns that value to the pixel. Black pixels were not evaluated against the user-specified color value, but automatically converted to white.

The program then enters another loop to segment the pink, red and blue pixels. The loop takes the image that now only contains pink, red, blue and white pixels and creates three different images, with each color assigned to an image (see figure 7). The three images are then converted to black and white using the MATLAB function `im2bw`. After they are converted to black and white, the function `bwarea` is used. `bwarea` estimates is the area of the objects. The scalar value given is a rough estimate of the total number of pixels in the image. The

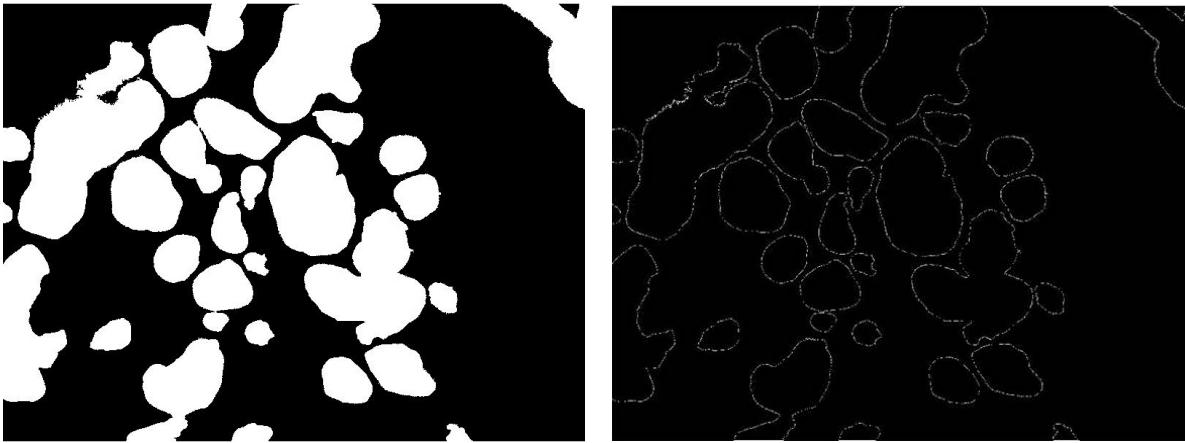


FIGURE 8. Total area as segmented by MATLAB (left); total perimeter of an image as measured by MATLAB (right).

program then combines the segmented images using the function `imadd`. It must be used twice in order to combine all three images, since it can only combine two images at once. We then use `imfill(image_name, 'holes')`, with the `holes` option, to fill dark areas of pixels surrounded by light pixels. `bwarea` and `bwperim` were then run on the images, which return the total area and total perimeter, respectively. `bwperim` produces an image of only border pixels, as seen in figure 8. This warrants another use of `bwarea` on these outline images to measure the perimeter. `bwarea` simply counts the number of white pixels present in each image.

2.5.2. Photoshop Selection Measurement Tool. We used Photoshop CS5 to measure the villi area and perimeter. This application also automatically calculated the circularity using the area and perimeter of each villi[1]. The specific process of measurement is as follows:

- (1) Open desired image.
- (2) Click “Window” > “Measurement Log.”
- (3) Select Lasso tool.
- (4) Trace the border of a villus that is not touching the edge.
- (5) In the “MEASUREMENT LOG” toolbar, click “Record Measurement.” A line with the desired measurement for the selection’s area, perimeter, and circularity are calculated.
- (6) Select Paintbrush or Pencil tool. Mark the measured villus.
- (7) Repeat steps 3-6 until all non-edge villi are marked.
- (8) In the “MEASUREMENT LOG” toolbar, click the icon to “Select all Measurements.”
- (9) In the “MEASUREMENT LOG” toolbar, click the icon to “Export Data.”
- (10) Save measurements. They will be saved in a tab-delimited text file.
- (11) In the “MEASUREMENT LOG” toolbar, click the icon to “Delete Measurements.”
- (12) Open next image.

The text files can then be imported into Microsoft Excel.

Although there is no absolute measure of accuracy for the process, accuracy was doubtless much improved overall by the use of a digital pen and tablet, rather than mouse. Despite this, larger villi presented a problem; the selection must be made without lifting the pen, so villi greater than a certain size were measured with a smaller image zoom. This made the pixels on the border more difficult to see and is possibly a source of error, in addition to simple human inaccuracy.

As it is probably apparent from the profuse number of steps and level of repetition necessary to measure an entire slide, the Selection Tool method is a lengthy process for even a small sample size such as this. For 64 images comprising subsections of full-size images, it took roughly 20

hours to complete all measurements. Although invaluable for gathering our data, this process is probably not a practical method for a larger sample size.

2.5.3. *Blood Vessels.* In order to count the blood vessels in each slide, Photoshop’s Count Tool was used. Count Tool is located under Analysis in the Tools panel. Photoshop’s Count Tool can be used to manually count items on an image or to reference or label items [1]. One must click on the items to be counted and Photoshop then keeps track of the number of clicks. After clicking on an item the sequential count number will appear next to the item. The number of items is recorded on the Measurement Log by clicking on “Record Measurements,” which is located right above the Measurement Log. One may change the marker size, the color, and the label size if desired. One is also allowed to clear any count measurements if it needs to be reset for any reason [1].

In some of the histology slides, the blood vessels had already been traced in green. By using the Tool Count, we simply clicked on all the traced blood vessels and once all the blood vessels had been identified we recorded the number of clicks, which is the same as the number of blood vessels. The Measurement Log contains the name of the slide (the lab identification number), date and time, the source (Tool Count in this case), and the count. The Measurement Log was then exported to an Excel file so that the data could then be easily analyzed. For the remaining images, we were to identify all the blood vessels and proceed as normal. Only the blood vessels in “whole” villi were considered, so any villus overlapping the image edge were disregarded.

The number of blood vessels on each image ranged from 28 to 228. Since full-size slides were divided into separate images, if a slide had more images than others, then it had disproportionately more counted blood vessels. For example, the total number of blood vessels in placenta 1479 was the sum of the counts of image no. 1479_1 + 1479_2 + 1479_3. For this reason, the average number of blood vessels per slide was calculated and was used for analysis.

3. RESULTS

There were numerous sources of possible error present in the measurement sets we obtained through our research methods. We account for all errors first, in order to explain later reasoning. We also attempt to mitigate the error magnitudes by eliminating certain variables, transforming measurements to calculate data sets, and conservatively analyzing variables.

3.1. Sources of Error.

- The analytical significance one can attribute to such a small data set is very low. Outliers have greater impact on trends, which are poorly indicated to begin with due to sparse distribution.
- The majority of the circumstances under which placentas and slides were prepared, and the methodology governing data collection from the UNC Excel file are unknown to us. We therefore cannot advise attributing undue statistical significance to our ad-hoc analysis of data where the hypotheses of the study are not known.
- None of our research group are trained histologists. Although we identified the anatomical structures to the best of our ability, referring to and learning from excellent medical texts, it is possible that we identified an object as a blood vessel which was in actuality a fibrin deposit, edema, tears accumulated under the torsion of slicing preparations; it is then also probable that we dismissed blood vessels, mistaking them for the aforementioned. Similarly, for the Photoshop Selection tool method, villi which were separate (connected) may have been traced as connected (separate). The best of care was taken to distinguish this by the use of trophoblastic cover, but many visually ambiguous cases presented themselves.
- Using the human hand as a tracing tool is naturally inaccurate for the Photoshop Selection tool method. Although the researcher who performed all the Photoshop Selection

tool measurements has lifelong drawing and illustration experience, including three years of digital pen-and-tablet illustration experience, simply the blood pressure of a heart-beat can shake the hand, subtracting or adding pixels in the process. Add to this the numerous factors affecting human motor dexterity or eye-hand coordination (caffeine intake, eye-fatigue, etc.) and the list grows relatively more exhaustive.

- The fact that villi with edges extending off the image border were excluded in the Photoshop Selection tool method and the blood vessel counting method may skew the data in favor of a larger distribution of smaller area villi. The very reason a villi may extend off the edge is that it is relatively large.
- The fact that villi with edges extending off the image border were included in the MATLAB measurement method skews the data toward larger total area measurements than would be obtained with the Photoshop Selection tool method. Therefore it prevents comparison of area measurements between those data sets (automated vs. by-hand).
- The varying number of images representing one placenta:
 - (1) Statistically favors placentas with several images for distributions of blood vessel counts and circularity. Since the average of these measurements over a placenta's image set was taken to represent the placenta (rationale to follow), there was a greater sample size of total measurements. Thus the measurement distributions of placentas with fewer images are vulnerable to outliers and their distribution patterns are weaker.
 - (2) Statistically may affect total area measurements from the MATLAB measured image sets.
- We received no information on the length scales of any images; hence all units of measure and data are given in pixels or adjusted to be dimensionless (rationale and method to follow). We also have no data on whether separate quadrant images are zoomed out or in relative to their set, and cannot completely trust the assumption that they are the same resolution.
- In the MATLAB measurement method, for the MATLAB-cleaned images, some maternal blood cells connected to villi by a sufficiently large number of pixels were recognized as being part of the villi; therefore maternal blood cells may have been included in this image set's total area measurement, adding to the area magnitude.
- For the MATLAB-cleaned images, shapes too small to meet a threshold, including blood cells and smaller villi, were deleted from the image in the cleaning process. The intention was to delete blood cells that were a pink similar to the color of the villi, but some villi also fell below threshold.
- An aggregate of the above two problems, some dense clusters of red blood cells that were a pink close to the color of the villi were both recognized as a single object, and was too large to be deleted by size thresholding. These cases erroneously increase the total area magnitudes.
- Some images were notably difficult to clean and segment automatically, since the placentas were insufficiently drained of blood, or because inflammation or pathological congestion was present. Consequently, there were a relatively substantial amount of images possessing a mixture of problems related to blood clusters, blood/villi confusion, and villi surrounding blood clusters; they have the associated errors as well.

3.2. Eliminated Variables. Initially we wished to calculate circularity using automatically segmented and measured images, but several problems mentioned above prevented this from being reliable. Most notable is the simplicity of the MATLAB measurement algorithm; subsequently it could not perform well on many slides with inadequate blood drainage. The splitting of the images and inclusion of the edge villi also prevented the perimeter data from MATLAB-measured images from being viable for any medical analysis. Consequently, due to the large

Cleaning Method	Measurement Method	Measurements	Data Transformation	Variable
P.S.	MATLAB	Total V.A.	Total V.A./Image A.	Fractional A.
P.S.	MATLAB	Total V.P.	None	None
MATLAB	MATLAB	Total V.A.	Total V.A./Image A.	Fractional A.
MATLAB	MATLAB	Total V.P.	None	None
None	P.S. Sel. Tool	A.	Sum over images	A.
None	P.S. Sel. Tool	P.	Sum over images	P.
None	P.S. Sel. Tool	Circ.	Avg. over images	Avg. Circ.
None	P.S. Count Tool	# of BVs	Avg. over images	BV Density

TABLE 1. How the measurements were transformed into variables. Abbreviations: P.S. = Photoshop, A. = Area, P. = Perimeter, V. = Villi, Sel. = Selection, BV = Blood Vessel, Avg. = Average, Circ. = Circularity.

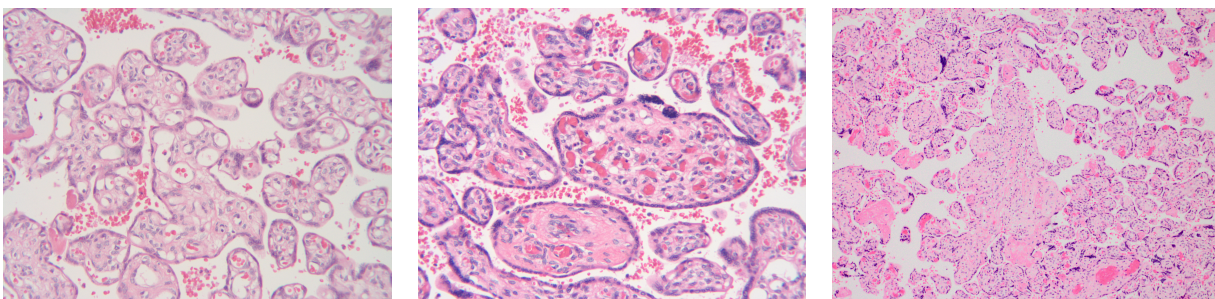


FIGURE 9. Three cases of differing villi shapes. From left to right: #1632, #1657, #2523.

differences in what parts of the images were measured and how, there is no objective way to compare the hand-traced Photoshop data to either automatically segmented and measured data sets.

We have, however, provided comparison of MATLAB and Photoshop cleaning methods to each other, since they were both measured using the same algorithm. To provide a variety of perspective on the performance of the cleaning methods, we have included perimeter in the Error Analysis section.

3.3. Data Transformation. Table 1 describes the transformation of data to compensate for some of the error or inaccuracy present in the data set or measurement methods.

3.4. Distribution of Data Sets.

3.4.1. Hand-Traced Data Comparisons Among Placentas. The hand-traced data, collected using the Photoshop Selection tool method, offers a somewhat reliable glimpse into the shapes of villi present in a cross-sectional slide of a placenta. Three placentas were chosen for analysis and comparison (see Figure 9): Placentas #1632, #1657, and #2523. With the best of our knowledge and ability learned from [3, 4], we selected #1632 to represent the Mature case of averagely sized and shaped villi; #1657 was selected to represent an Immature case, with villi of relatively larger shape; and #2523 represented the Hypermature case, with many relatively small villi [3].

See figures 9, 10, 11, 12. Each unit in the distributions represents the measurement of one villus. Thus the number atop each graph’s bars represents the number of villi falling into the indicated range (or “bin”) for that placenta. Since by nature the Photoshop method recorded one measurement per villus, for the distribution we included all measurements for one placenta; that is, all villi in all images representing the placenta. Each of these placenta had several images representing it. Note: the bin ranges for the distributions are different for each placenta.

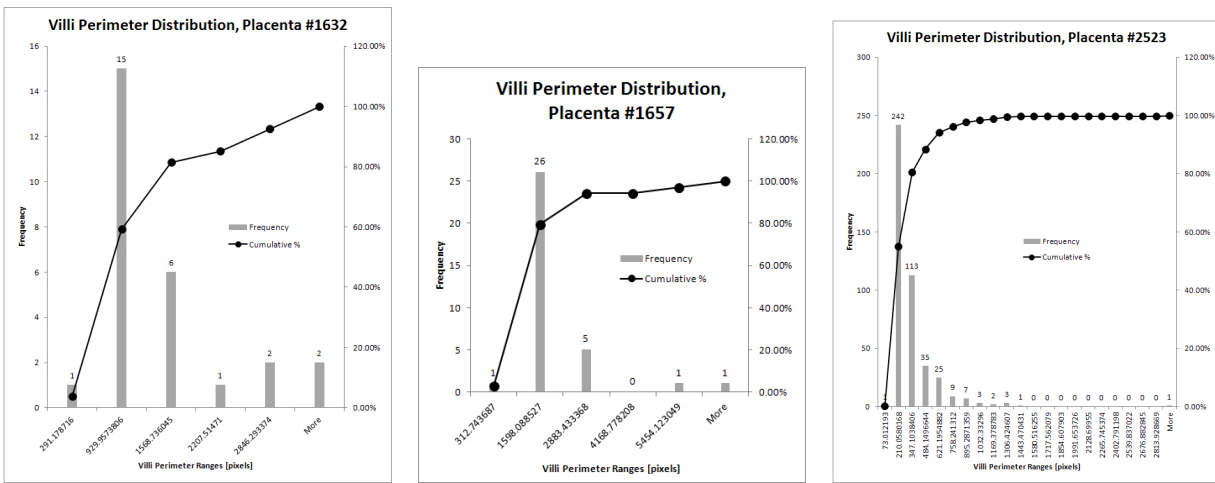


FIGURE 10. Perimeter distributions for the Photoshop Selection tool measurements of placentas #1632, #1657, and #2523.

Perimeter		
	50%	95%
#1632	930	2847
#1657	313-1600	2889
#2523	211	622

TABLE 2. Benchmarks for “half” (50%) and “most” (95%) cumulative distribution of perimeter size.

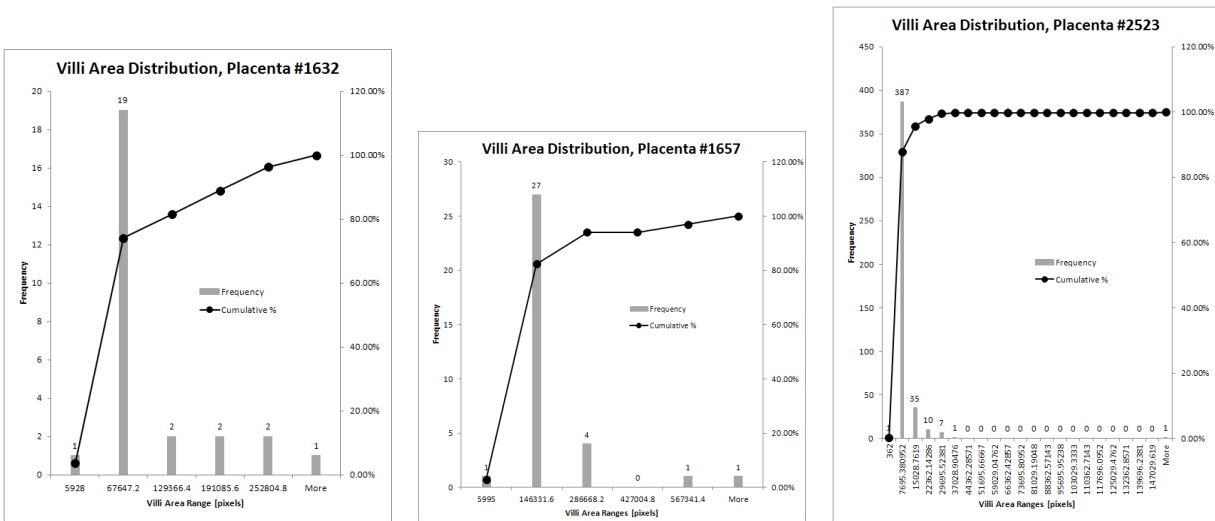


FIGURE 11. Area distributions for the Photoshop Selection tool measurements of placentas #1632, #1657, and #2523.

The cumulative percentage lines show at each range what percentage of the villi are at a pixel value below that range. In comparing placentas this becomes a useful tool, since we can see what perimeter, size, and proportions that half (50%) and most (95%) villi are. These results are summarized in each of the tables 2, 3, 4, where all values given are in units of pixels.

Thus #1632 and #1657 are roughly the same in villi perimeter distribution, whereas the perimeters of #2523 are smaller on average.

Area		
	50%	95%
#1632	5928-67647	252805
#1657	5995-146332	286663
#2523	362-7696	15029

TABLE 3. Benchmarks for “half” (50%) and “most” (95%) cumulative distribution of area size.

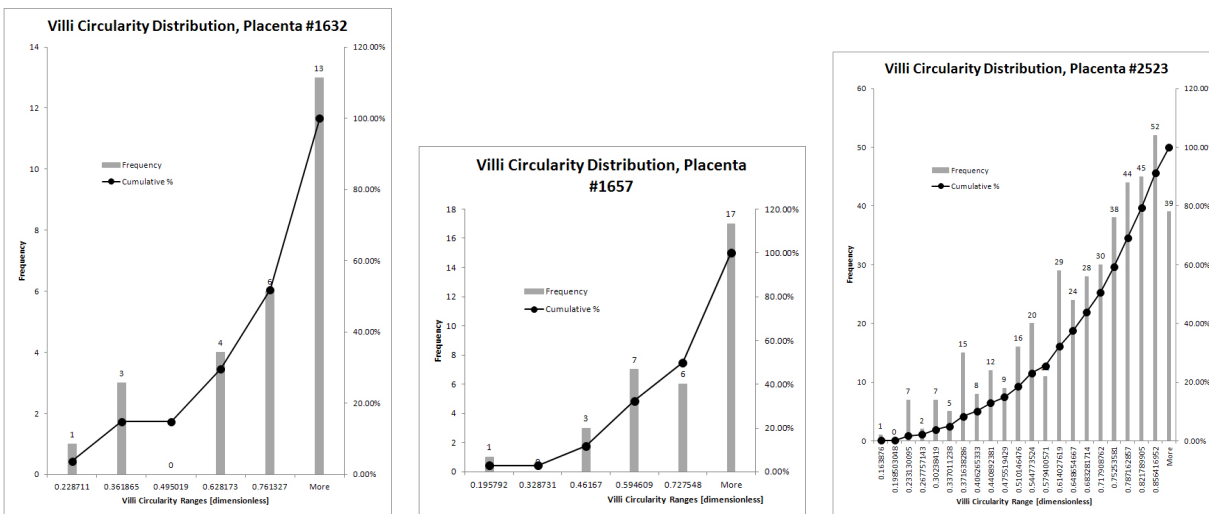


FIGURE 12. Circularity distributions for the Photoshop Selection tool measurements of placentas #1632, #1657, and #2523.

Circularity		
	50%	95%
#1632	0.761327	> 0.762
#1657	0.727548	> 0.728
#2523	0.717909	> 0.857

TABLE 4. Benchmarks for “half” (50%) and “most” (95%) cumulative distribution of circularity value.

Similar to the case with perimeter, the area distributions of the villi of #1632 and #1657 at 95% fall into roughly the same range. However at 50%, there is a difference in the upper end of the range for these placenta; the range given for #1657 is roughly double the pixel range for #1632. It is not clear if this necessarily implies that half the villi in #1657 are larger, as the bins were created at regular intervals, and the sizes could possibly be identical despite the bin labels. #2523 has markedly smaller villi area than the other two placentas.

Whereas the perimeter and area distributions followed roughly the same cumulative percentage line segment pattern, resembling perhaps a logarithmic function, that of the circularity resembles an exponential curve. For circularity, the greatest number of villi in each graph are at the higher end of the bin range, above a circularity of 0.762, 0.728, and 0.857, for placentas #1632, #1657, and #2523, respectively. The percent differences between each value pair are 2.3%, 5.9%, and 8.1%. This may confirm the noticeable visual differences between the images in Figure 9, especially confirming #2523 as relatively more mature than #1657.

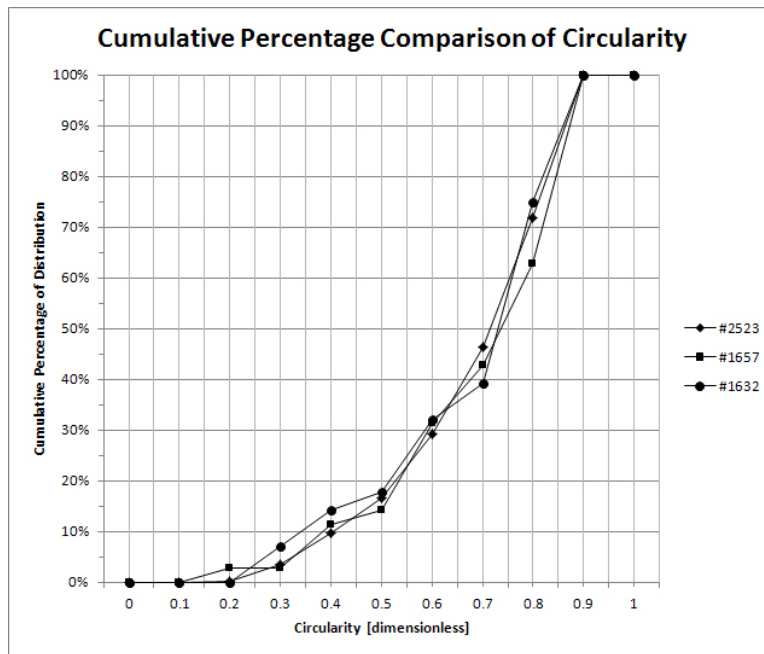


FIGURE 13. Comparison of the cumulative percentages of circularity for placentas #2523, #1657, #1632. Note the similar pattern the line segments follow.

The patterns of cumulative distribution lines create almost the same shape in each type of measurement, despite the differences in size and number of shapes traced in total. Below is a comparison chart between the cumulative percent distributions of circularity for the three placentas. Since circularity is a proportion which occupies a normalized numerical range, as opposed to a measurement like area which is patently different, it gives a way to compare the images despite dimensional measurements.

The similarity of the pattern of cumulative percentage is striking. Even though an 8.1% difference was detected between the circularity proportion of hypermature #2523 and immature #1657, figure 13 asserts that the image pair is still 91.9% similar.

It is difficult to conclude very statistically strong results from this analysis, for several reasons. In the absence of information about the microscopic resolution of the images, we have (poorly) assumed that each image is at the same zoom. If this were not the case, the differences in villi measurements we detected might not necessarily have to be disregarded; if we did have an image with a smaller zoom or smaller villi, there is not a fundamental difference in area or perimeter, only pixel measurement accuracy. It is more difficult to say whether zoom can distort the circularity. The need to disregard zoomed images may be indicated by the relative similarity of circularities; Further study on this issue would have to be conducted for confirmation, comparing zoomed and hypermature images.

It is also a problem inherent in the definition of hypermaturity to have a greater number of villi, whereas in immature cases the villi are far fewer and larger. For the same zoom value, this results in a greater “samples size” for a hypermature placenta image, and hence a smoother distribution. By extension, large villi in immature or mature images that are split from a full-size image tend to have more villi that can breach image edges and be disregarded.

Despite the problems, we believe circularity can be indicated as a possible measurement in future studies with a better data set. First, a diagnosis of chronic villitis, slightly preterm delivery (257 days), and low birth weight (1824g) were present in case #2523, and analysis detected a distribution of higher circularity values. This is consistent with what was expected from a hypermature placenta (refer back to figure 1), namely a relatively larger amount of terminal villi. This type of villus tends to be small, round and “grapelike” [3]. We confirm that

we have detected a relatively large quantity of small, relatively round shapes for a hypermature case.

In contrast, cases #1632 and #1657 were from infants of normal-weight and gestational age at delivery, without any glaring pathology present in data. Their villi also appear similar to each other, which was mostly confirmed by measurement, yet different compared to #2523, again confirmed by analysis of the distributions.

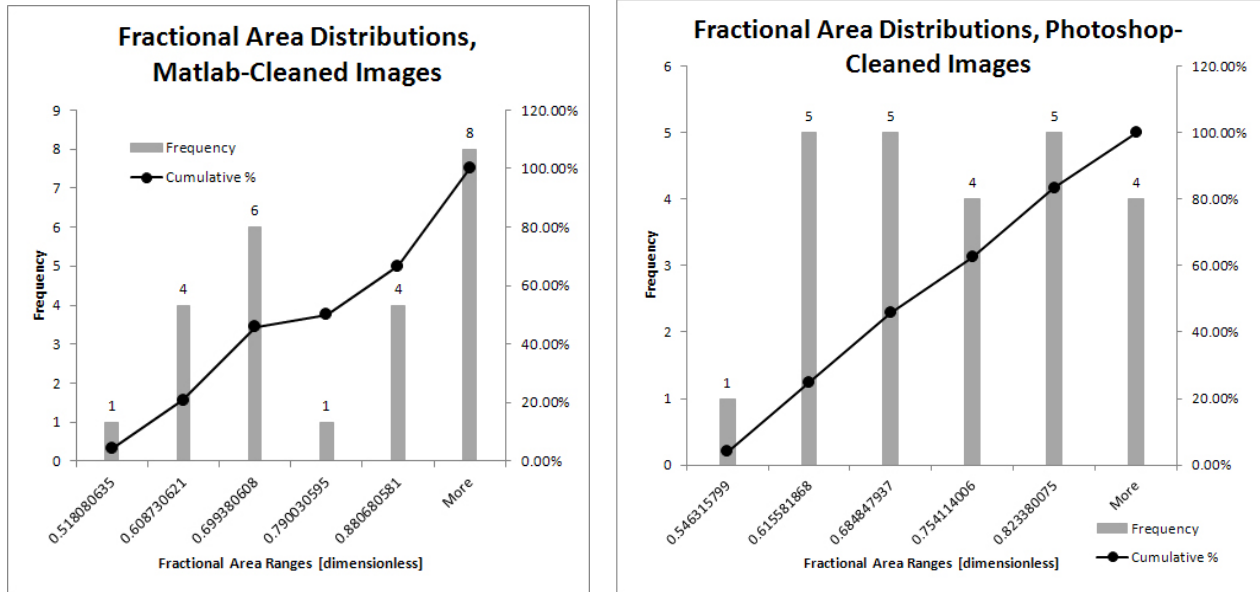


FIGURE 14. Histogram distributions of the fractional areas (see section 3.3) of the images cleaned in MATLAB (left) and Photoshop CS5 (right). Cumulative percentages are displayed.

3.5. Comparison of Fractional Area between Photoshop-Cleaned and MATLAB-Cleaned Images. As discussed in section 3.3, we needed to find a way to use the areas obtained by automatic measurement in MATLAB to represent an aspect of placental maturity. Since we previously detected a difference in villi area distribution between extreme cases, we compare the distribution of fractional area for all placentas.

The fractional area is a dimensionless measure of how much of a slide is occupied by villi. Therefore it gives insight into both how well each algorithm performed, and in the case of the Photoshop-cleaned images, provides a measure of fractional area which is better-suited for analysis. The latter analysis for placentas is performed in this section; see the former evaluation of algorithm vs. software methods in section 3.7, Error Analysis. The distinction between the area comparisons looked at in the previous section are per *placenta*; the section deals with total area measurements per *image*.

In the Photoshop histogram from figure 14, the cumulative percentage lines create a nearly straight line, and there is a roughly equal frequency across the entire range. It is difficult to gauge what an even distribution is supposed to mean, in terms of clinical significance. From available data, the placenta with the lowest fractional area (0.518) was #2723, from a normal-weight (3278g), term (39 weeks) infant; the highest fractional area of 0.971 from case #3042 was from a slightly high birth weight infant (4006g) which was also aged 39 weeks at delivery. A slightly preterm (258 days or barely under 37 weeks) case presented a fractional area of 0.829, and an underweight (1904g) but term delivery case presented with fractional area 0.650. In figure 15, there is erratic spread of data points and low R^2 values.

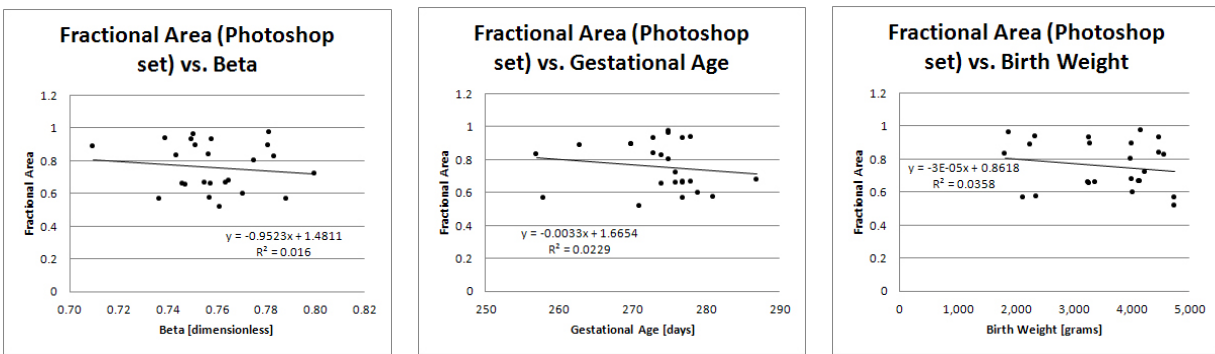


FIGURE 15. Linear regression using the Fractional Area variable versus medical variables.

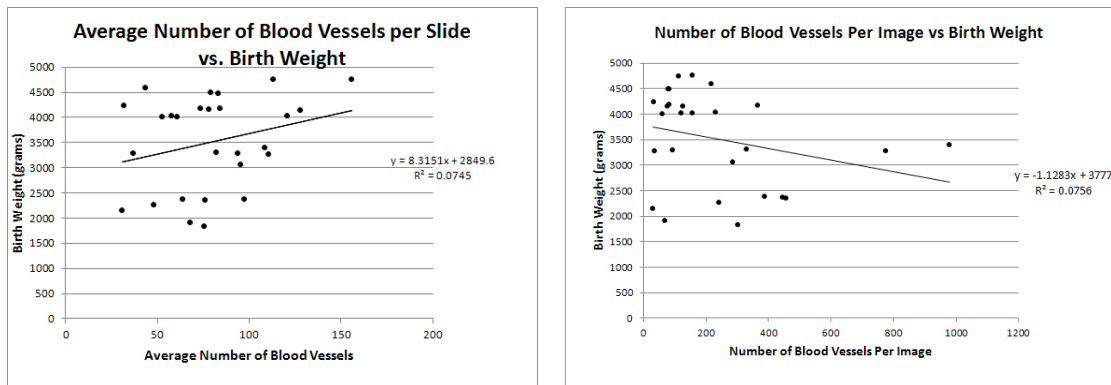


FIGURE 16. (a) Scatterplot of the average number of blood vessels per slide vs. birth weight. (b) Scatterplot of the total number of blood vessels per image vs. birth weight.

Therefore the fractional area is not a reliable measurement for use in diagnosis; however, rather than the admittedly low sample size of 27, some images could not be processed. Three were left out of the fractional area histograms (see bar labels in figure 14). The data set being small may have factored in to fractional area having little correlation with medical data. However, the fractional area is still vulnerable to the microscopic resolutions of the images.

3.6. Relationships between Medical Variables and Blood Vessel Measurements.

3.6.1. Linear Regression. In order to determine if there was a trend between the average number of blood vessels per slide and the birth weight of the newborns tested, the regression line and its R^2 value were calculated using excel. Figure 16 (a) is the scatter plot corresponding to the relationship between the average number of blood vessels (dependant variable) and the birth weight in grams (independent variable). Although the calculated regression line ($y = 8.3151x + 2849.6$) appears to fairly fit the data, the coefficient of determination R^2 was only 0.0745, implying that approximately seven percent of the variation in the birth weight can be explained by the average number of blood vessels per slide. Because this regression model did not fit the data well, we went on to comparing the total number of blood vessels in a slide and the birth weight (grams). This resulted in a regression line of $y = -1.1283x + 3777$ with $R^2 = 0.0756$. By using the total number of blood vessels, instead of the average number per slide, the regression line indicates that as the number of blood vessels increases the birth weight decreases, whereas by using the average number of blood vessels the slope of the regression line was positive (Figure 16 (b)).

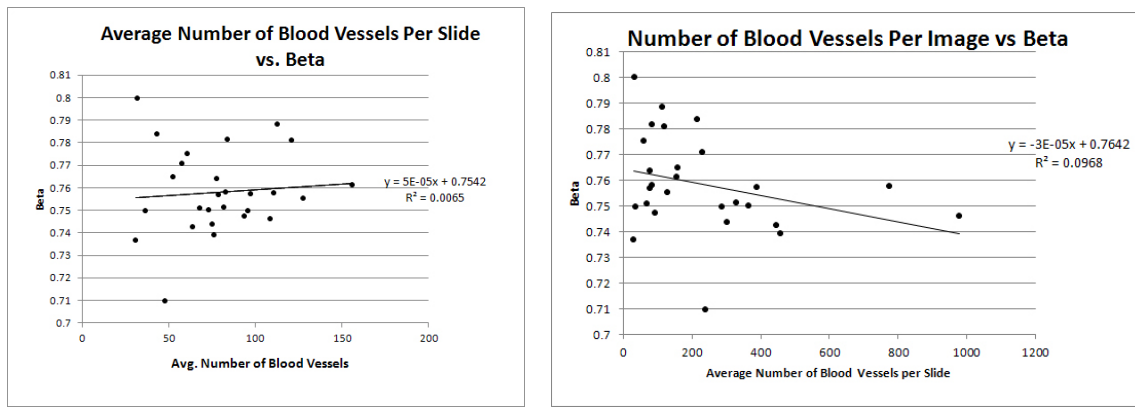


FIGURE 17. (a) Scatterplot of the average number of blood vessels per slide vs. β . (b) Scatterplot of the total number of blood vessels per image vs. β .

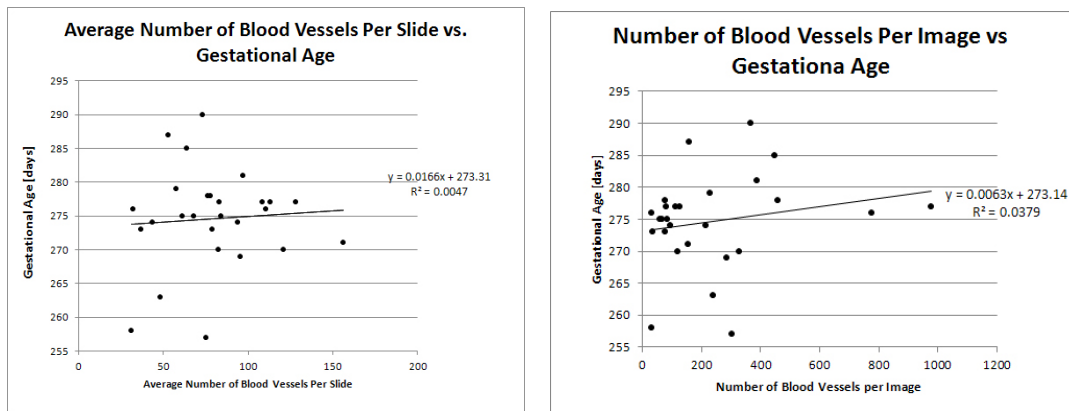


FIGURE 18. (a) Scatterplot of the average number of blood vessels per slide vs. gestational age. (b) Scatterplot of the total number of blood vessels per image vs. gestational age.

The average number of blood vessels was compared to β ($\approx \log \text{PM}/\log \text{FM}$) in order to determine if a trend could be detected [6]. Figure 17 (a) shows the scatter plot of these two variables. The regression line ($y = 5e^{-05}x + 0.7542$) and R^2 value (0.0065) indicates that outcomes will not be well predicted by this model. These results were compared to the total number of blood vessels vs. β and its regression line and R^2 value were determined (Figure 17 (b)). Although the value $R^2 = 0.0968$ improved, the regression line ($y = -3E-05x + 0.7642$) is still not a good model for the data.

Because we predicted the number of blood vessels was related to the gestational age of a fetus, the average number of blood vessels per slide was compared to the gestational age. Figure 18 show the scatter plot of the average number of blood vessels (independent variable) and the gestation age in days (dependant variable). For this pair of data the regression line is $y = 0.0166x + 273.31$ and $R^2 = 0.0047$. This model also shows a poor trend, so once again the total number of blood vessels was graphed against the gestational age and a scatter plot was created. Figure 18 shows that the new regression line is $y = 0.0063x + 273.14$ with an improved value of $R^2 = 0.0379$. Although R^2 improved it is still too low to be considered significant.

Table 5 summarizes the our calculations comparing the average number of blood vessels per slide to birth weight, β , and gestational age. Table 6 also summarizes our calculations using the total number of blood vessels per image. We believe our results using the average number of blood vessels and the total numbers of blood vessels are so different because when an image

Variable	Linear Regression Line	R^2
Birth Weight	$y = 8.3151x + 2849.6$	$R^2 = 0.0745$
β	$y = 5e^{-05}x + 0.7542$	$R^2 = 0.0065$
Gestational Age	$y = 0.0166x + 273.31$	$R^2 = 0.0047$

TABLE 5. This table shows the regression line and R^2 value for the average number of blood vessels per slide and each of the variables on the table.

Variable	Linear Regression Line	R^2
Birth Weight	$y = -1.1283x + 3777$	$R^2 = 0.0756$
β	$y = -3e^{-05}x + 0.7642$	$R^2 = 0.0968$
Gestational Age	$y = -3e^{-05}x + 0.7642$	$R^2 = 0.0379$

TABLE 6. This table shows the regression line and R^2 value for the total number of blood vessels per image and each of the variables on the table.

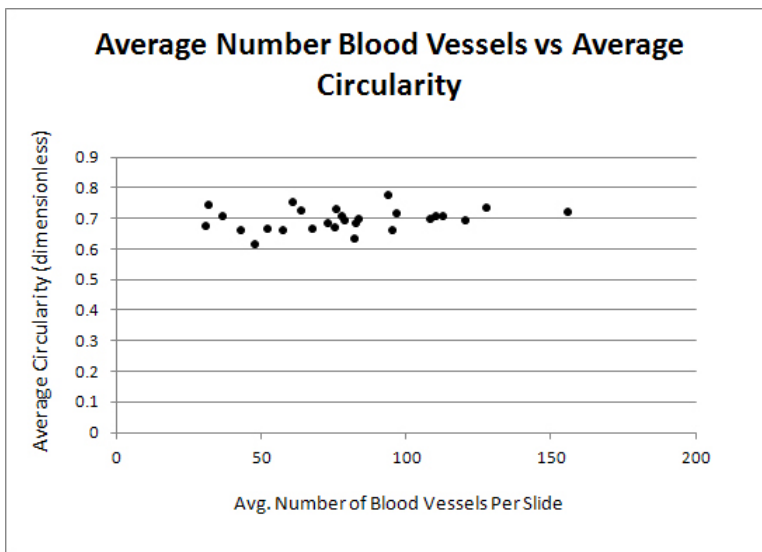


FIGURE 19. Scatter-plot of the average circularity for all villi in a given placenta vs. the average blood vessels per slide.

is split into quadrants the density of blood vessels may be different in each quadrant. Also, since partial villi were disregarded when counting vessels, when vessels of the slides were added together there are many villi that are not taken into account. Detecting a trend is also difficult because we are only comparing two factors at a time when in fact, the socioeconomic status, ethnicity, or number of previous birth of the mother might make a difference in our data [6, 7, 2]. Because we had a small sample size (27 images) the sampling error in our data is large compared to the sampling error if our sample size had been bigger.

3.6.2. *Correlations.* We also sought to find a correlation between the average number of blood vessels per slide and the circularity variable from the hand-traced data. Figure 19 shows the scatter plot for the average number of blood vessels per slide vs. average circularity.

By using Excel’s CORREL function, the correlation between these pairs of data were calculated. The correlation between the average number of blood vessels and average circularity was only 0.257858578 indicating there is a small degree of dependence between these two variables; this is still a relatively large correlation compared to other data trends we looked at with linear

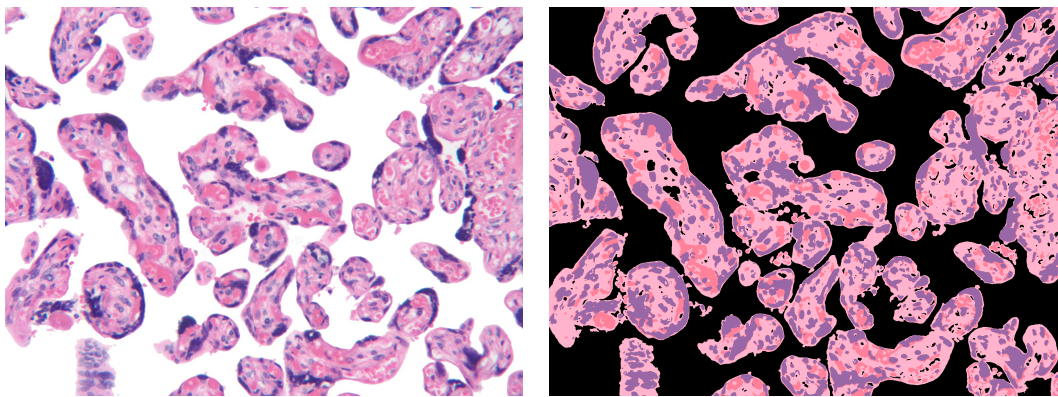


FIGURE 20. Comparison of Photoshop and MATLAB images of slide 1479_1.

Color	Red	Pink	Purple	Total	Perimeter
Min Error	87	868	0	32686	535
Max Error	1134717	114105	14684	1389373	24661

TABLE 7. Minimum and Maximum errors in comparing the MATLAB-cleaning algorithm and the Photoshop-cleaning method. All values are measured in pixels.

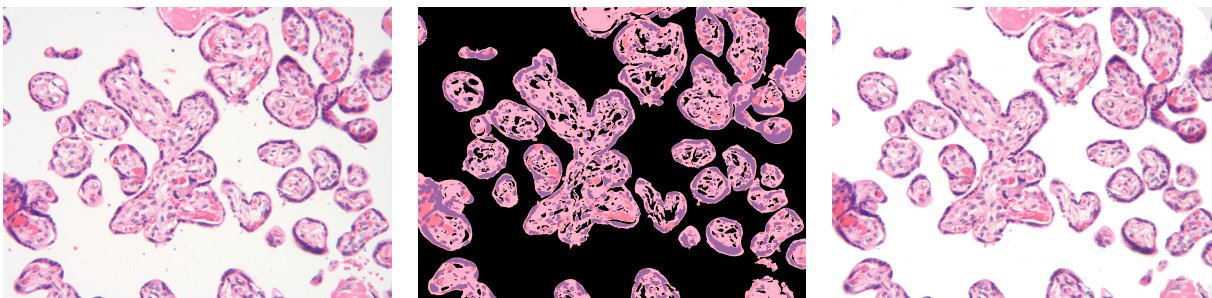


FIGURE 21. Images of a good segmentation result, #1737.

regression. The reason may be because human-powered methods are able to differentiate between joined villi and edge-villi. Because we had a small sample size, the outliers had a larger distorting effect on the correlation.

3.7. Comparison of MATLAB & Photoshop Image Cleaning Through Error Analysis. In order to make a comparison as to how accurate our MATLAB-cleaning algorithm was, we needed something to compare to. Using Photoshop, we were able to make a “perfectly” cleaned image and used this with the same color markers used in the algorithm to compare our results.

As seen in this Photoshop-cleaned image (figure 20) the results are similar to the one obtained via MATLAB beside it, but it is not a perfect match. The discrepancies in the image comes from the blood we want to remove being too close to the villi, or being in one large group. Since the algorithm relies on color markers, if an single cell near a villus is connected to it by even one pixel, then the algorithm thinks that the cells are part of the villi and will not remove it. Also If the cells we want to remove are in one big group, because of the way we remove small objects, the code will not recognize that group as something we wish to remove; it will interpret the cluster as a large object and fail to delete it. Table 7 shows the minimum and maximum errors for the different sets of images.

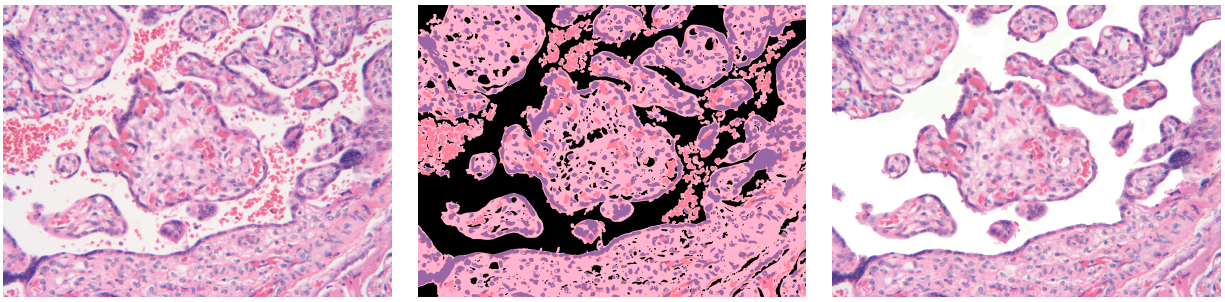


FIGURE 22. Images of a poor segmentation result, #1657.

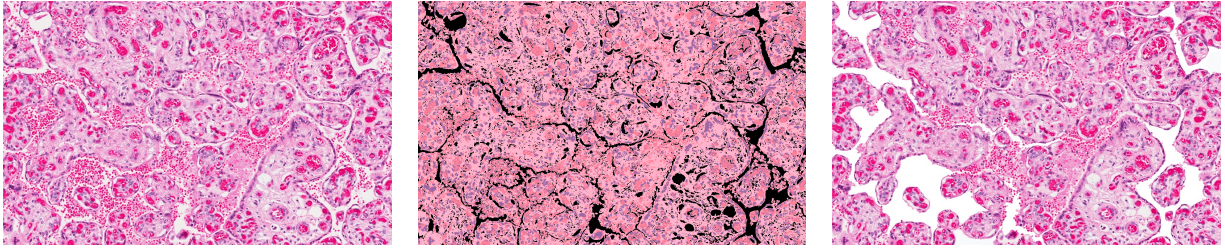


FIGURE 23. Images of a very poor segmentation result, #2870.

Original images that do not have many maternal blood cells achieve very good results (figure 21), while images that have many cells in the intervillous space (see figure 22) are not cleaned very much (figure 23).

Although figure 20 shows that the two images are similar, the table implies that there is a big difference between minimum and maximum error. For consistent results, the difference between the maximum and minimum error must be minimized. Figures 21, 22, and 23 show the cleaning results for placentas #1737, #1657, and #2870 (respectively) between the original, MATLAB, and Photoshop images (from left to right).

4. FUTURE WORK

Despite its usefulness for gathering data, the hand-tracing and hand-counting methods for gathering circularity and blood vessel data is not a practical method. They require too much time of a skilled histologist in order to assure accuracy. However, our simplistic algorithm also performed segmentation poorly and was not sophisticated enough to return useful variables.

If more time had been available, our group would have searched for a way to automatically count the blood vessels in each slide using MATLAB. We would have then been able to compare our MATLAB results to our Photoshop results. This would have taken a long time since blood vessels are either pink or white on the slides and distinguishing between different shades of pink is a difficult task.

Although our segmentation results were good for some images, it is still not a viable option for future studies, mainly because our algorithm relies of predetermined color markers. All of the automated algorithms we tried also failed to provide reliable results. Another method we would like to try would be to develop a way to get the individual color markers per image; this may obtain more accurate results. We would also like to find a more consistent way of removing large clusters of maternal blood cells that our code could not remove.

We also limited ourselves to using MATLAB for this project, and we would like to try other programs that are available such as C++ and Java. Although these codes can be rewritten to match any language, we feel that using a combination of different programs can improve our results.

Even with the ideal data set of full-size placental histology images for every mother-infant-placenta variable set (1277 in total), along with more information about the conditions of the study and slide resolutions, the problem of creating an algorithm that is sufficiently sophisticated and accurate is daunting. Supposing we had more time, resources, and images, it is possible to, if not develop software to do so, use Photoshop or another image analysis software suite, to take a definitive set of shape factor measurements. The larger sample set would facilitate more sophisticated exploratory data analysis techniques such as principal component analysis, cluster analysis, or multivariate regression to uncover placental maturity and medical variables. With the hypotheses this data mining suggests, other experiments could be designed that could independently confirm the suspicions using a different sample set.

Ultimately, the idea behind studying placental maturity is to construct a quantitative measure of placental development which has predictive power for an infant entering the world. The degree to which the placenta's maturity is predictive for the long-term health of the infant is established [3]. Benchmarks for fetal physical and neurological development are already established and are in more frequent use. In order to create a placental maturity reference range with quantitative categorical distinctions, large amounts of data must be collected, then processed. A sufficiently sophisticated piece of software might accomplish the data collection, then be used again on a fresh slide, in order to compare it to the collected data.

5. CONCLUSIONS

We cannot extrapolate any significant generalizations using a small sample set and measurement methods with compounding inaccuracy to the population at large. We can, however, report mild success in connecting images that have been given a categorical value of mature, hypermature, or immature, with relative differences in the distribution of the geometric properties of villi. Of all analysis performed on data, this provides the most promise in regards to the medical value of the information. However, a less-work intensive approach to measuring villi must be developed before it is viable to study their shape factors with large sample size or image size. Specifically, a more complex, robust collection of algorithms must be developed to differentiate between villi, account for all instances of fibrin deposits, blood vessels and maternal blood cells, and measure the shape factors desired.

The linear regression models that were created comparing the average number of blood vessels to the birth weight, β , and gestational age resulted in very low R^2 values. This indicates that a very low percentage of the variation in the three variables we tested could be explained by average number of blood vessels per slide. The model that has the highest R^2 value was the linear regression line for the birth weight, but it is still not significant enough to indicate a trend. Thus we detected no linear trends in the data set between number of blood vessels and any medical variable.

The correlation between the average number of blood vessels and the average circularity per placenta was low but not completely insignificant, a relationship may possibly exist.

Due to the complexity of some of the images, the segmentation and measurement results from MATLAB vary significantly from that of Photoshop. And although we had some positive results, the disadvantages of the algorithm outweighed the advantages. Predetermined color markers based on an average of all images cannot be used as a standard value to compare to simply because the variation in color between images is too great. The method in which maternal blood cells are removed is crude. Simply removing small areas would not work for hypermature placentas that may contain small villi. And finally, this method cannot remove cells that are attached to the walls of the villi simply because there is no way to differentiate between the two, unlike Photoshop, where the user can, by hand, remove them.

The main advantages of using Photoshop cleaning are being able to remove elements with human judgment and the ability to separate the villi. These result in a better measurement of

the area and perimeter than the MATLAB-segmented images. As previously mentioned with regards to the Selection tool method, human-operated processes are time consuming.

6. ACKNOWLEDGEMENTS

This research project could not have been possible without the help of Dr. Salafia who provided the histology slides and the medical and image information corresponding to each slide. In addition, Dr. Salafia went out of her way and invested time going over the work we have done, such as our presentation.

Profuse thanks are due to Dr. Chang, whose experimental industrial mathematics seminar class structure facilitated a stimulating research environment. We are grateful that Dr. Chang provided both challenge to our intellect and aid in our endeavours, as well as insightful criticism, and the important guidance needed to bring this project as far as it has come.

A special recognition needs to be given to our classmates and the other interested students for expressing their constructive criticism and suggestions, which inspired us to develop stronger ideas and methods.

We would also like to give thanks to the assistance of the CSU Long Beach librarians Catherine Outten, Eileen Waikiji, and Hema Ramachandran, who provided valuable research resources and their equally valuable time.

APPENDIX A. MATLAB SOURCE CODE

A.1. Image Segmentation: “Modeling_Project”.

```
%Segmentation of the 4 primary colors
pathname = '\histology images\';
dirlist = dir([pathname '*.tif']);
for x = 1:length(dirlist)
    images{x} = imread([pathname dirlist(x).name]);
end
for i = 1:length(images{x})
    name = dirlist(i).name;
    str = sprintf('Processing file %s',name);
    disp(str)
    %Get images and color images ready
    I = images{i};
    I = im2double(I);
    Iorig = I;
    Igrayorig = rgb2gray(I);
    [n,m] = size(Igrayorig);

    %Primary color markers
    red = [1 .5 .6];
    pink = [1 .7 .8];
    blue = [.6 .4 .65];
    white = [1 1 1];
    black = [0 0 0];
    %Color simplification and segmentation
    for k = 1:n
        for j = 1:m
            J = [I(k,j,1) I(k,j,2) I(k,j,3)];
            r = norm(J - red);
            p = norm(J - pink);
```

```

    b = norm(J - blue);
    w = norm(J - white);
    color = [r p b w];
    change = min(color);
    if change == r
        I(k,j,:) = red;
        Ired(k,j,:) = red;
    end
    if change == p
        I(k,j,:) = pink;
        Ipink(k,j,:) = pink;
    end
    if change == b
        I(k,j,:) = blue;
        Iblue(k,j,:) = blue;
    end
    if change == w
        I(k,j,:) = white;
    end
end
end

Iseg = I;
Igray = rgb2gray(Iseg);
%Grayscale image to black and white
for k = 1:n
    for j = 1:m
        if Igray(k,j) == 1
            Igray(k,j) = 0;
        else
            Igray(k,j) = 1;
        end
    end
end
end
%Removing small areas
Iadd = imadd(Ipink,Iblue);
Iadd = imadd(Iadd,Ired);

for k = 1:n
    for j = 1:m
        if Iadd(k,j) == 0
            Iaddbw(k,j) = 0;
        else
            Iaddbw(k,j) = 1;
        end
    end
end
end

area = bwarea(Iaddbw);
area = int32(.004*area);
area = double(area);

```



```

Iadd2 = bwareaopen(Iadbbw,area);

%Creating Mask of cropped image
for k = 1:n
    for j = 1:m
        if Iadd2(k,j) == 0
            Iseg(k,j,:) = black;
        end
    end
end
%Segmentation of cropped image
for k = 1:n
    for j = 1:m
        J = [Iseg(k,j,1) Iseg(k,j,2) Iseg(k,j,3)];
        if norm(J-red) == 0
            Ired(k,j,:) = red;
        else
            Ired(k,j,:) = black;
        end
        if norm(J-pink) == 0
            Ipink(k,j,:) = pink;
        else
            Ipink(k,j,:) = black;
        end
        if norm(J-blue) == 0
            Iblue(k,j,:) = blue;
        else
            Iblue(k,j,:) = black;
        end
    end
end
end

%strname = sprintf('%s',name);
%imwrite(Iseg,strname,'bmp');
end

```

A.2. Image Measurement: “Find_Areas;”

```

clear all
%path to the directory of the images
dirname = './images2process';
imdir = dir(dirname);

%determines how many images to process
nI = length(imdir)-2;
D = cell(1,nI);
mI = nI+2;
x = 1;

%loop to process images to determine areas and perimeter for each image
for i = 3:mI
    fname = imdir(i).name;

```

```

disp([num2str(i-2) ' of ' num2str(nI) ' processing image ' fname ' ' ]);
%changes directory to where images are located
cd ./images2process
%reads image etc
I = imread(fname);
I = im2double(I);
Iorig = I;
Igray = rgb2gray(I);
[n,m] = size(Igray);
%preset colors
red = [1 .5 .6];
pink = [1 .7 .8];
blue = [.6 .4 .65];
white = [1 1 1];
black = [0 0 0];
%loop to determine pixel's colors
for i = 1:n
    for j = 1:m
        J = [I(i,j,1) I(i,j,2) I(i,j,3)];
        r = norm(J - red);
        p = norm(J - pink);
        b = norm(J - blue);
        w = norm(J - white);
        bk = norm(J - black);
        color = [r p b w bk];
        change = min(color);

        if change == r
            I(i,j,:) = red;
        end
        if change == p
            I(i,j,:) = pink;
        end
        if change == b
            I(i,j,:) = blue;
        end
        if change == bk
            I(i,j,:) = black;
        end
    end
end

Iseg = double(I);

%changes to previous directory
cd ../..

Ired = Iseg;
Ipink = Iseg;
Iblue = Iseg;
Iwhite = Iseg;

```

```

for i = 1:n
    for j = 1:m
        J = [Iseg(i,j,1) Iseg(i,j,2) Iseg(i,j,3)];
        if norm(J-pink) == 0
            Ipink(i,j,:) = white;
        else
            Ipink(i,j,:) = 0;
        end
        if norm(J-red) == 0
            Ired(i,j,:) = white;
        else
            Ired(i,j,:) = 0;
        end
        if norm(J-blue) == 0
            Iblue(i,j,:) = white;
        else
            Iblue(i,j,:) = 0;
        end
        if norm(J-white) == 0
            Iwhite(i,j,:) = white;
        else
            Iwhite(i,j,:) = white;
        end
    end
end

Ipink = im2bw(Ipink);
Ired = im2bw(Ired);
Iblue = im2bw(Iblue);

Pink = round(bwarea(Ipink));
Red = round(bwarea(Ired));
Blue = round(bwarea(Iblue));

TArea = imadd(Ipink, Ired);
Iblue = double(Iblue);
TArea = imadd(TArea, Iblue);

TArea2 = imfill(TArea,'holes');
TArea2 = bwareaopen(TArea2, 25);

%Tests if Total Area of villi equals Total Area of image if it does
%then function imfill() will not be used to calculate Total Area of
%villi. This will only be used for extreme cases where images are
%nearly completely pink, red and blue. We only had 2 out of 95
%of our images that needed imfill not to be used to calculate total
%area and perimeter
Test = bwarea(TArea2);
Iwhite = im2bw(Iwhite);
Awhite = round(bwarea(Iwhite));

```

```

if Test == Awhite
    TA = round(bwarea(TArea));
    P = bwperim(TArea);
else
    TA = round(bwarea(TArea2));
    P = bwperim(TArea2);
end

Perimeter = round(bwarea(P));
Pink = round(bwarea(Ipink));
Red = round(bwarea(Ired));
Blue = round(bwarea(Iblue));

M(x,1) = TA;
M(x,2) = Perimeter;
M(x,3) = Pink;
M(x,4) = Red;
M(x,5) = Blue;
x=x+1;

i = i+1;
end
%Displays Total Area, Perimeter, Pink area, Red area, Blue area
disp(M);
disp(' ');
disp('Total Area, Perimeter, Pink area, Red area, Blue area:');

```

REFERENCES

- [1] Adobe Systems Incorporated. *Performing a measurement (Photoshop Extended)*, 2011.
- [2] Morten Andersen, David Belangery, Radina Droumeva, Jenny Li, Gilbert Moss, and Gabriela Palau. Quantifying clinically significant features of placental histology images: a method. Pdf file, (11,000k), World Wide Web, 2008.
- [3] Kurt Benirschke and Peter Kaufmann. *Pathology of the Human Placenta*. Springer-Verlag, New York, New York, fourth edition, 2000.
- [4] Luiz Carlos Junqueira, Jose Uchoa Carneiro, and Alexander Contoulpolos. *Basic Histology*. Lange Medical Publications, Los Altos, California, ninth edition, 1998.
- [5] The MathWorks, Inc. *Image Processing Toolbox: User's Guide*, version 7.1 edition, September 2010.
- [6] C.M. Salafia and M. Yampolsky. Metabolic scaling law for fetus and placenta. *Placenta*, 30:468–471, 2009.
- [7] C.M. Salafia, M. Yampolsky, D.P. Misra, O. Shlakhter, D. Haas, and B. Eucker J. Thorp. Placental surface shape, function, and effects of maternal and fetal vascular pathology. *Placenta*, 31:958–962, 2010.
- [8] Bagon Shai. Shai bagon's matlab code, June 2007.
- [9] F. Stoz, R. A. Shuhmann, and B. Schebesta. The development of the placental villus during normal pregnancy: Morphometric data base. *Archives of Gynecology and Obstetrics*, 244:23–32, 1988.