The Automated Calculation of Human Ovarian Reserve

Tom Kelsey

[1]School of Computer Science, University of St Andrews, St Andrews,
Fife, KY16 9SX, UK
tom@cs.st-and.ac.uk

Abstract. We report on the use of AI techniques – including image analysis, feature recognition, machine learning and constraint satisfaction – applied to the problem of determining the population of primordial follicles in a human ovary. We highlight the importance of accurate population data for studies that aim to predict the age of menopause in healthy women, and for studies that quantify damage to ovarian tissue as a result of radio- or chemotherapy. We describe existing methods used to **estimate** primordial follicle populations, and review the strengths and weaknesses of these approaches. We present a novel methodology for counting **each and every** primordial follicle in an ovary, together with details of our implementation and empirical evaluation.

1 Introduction

The remaining reproductive lifespan of every female human is determined by the population of primordial follicles contained in her ovaries. This population can not be measured *in vivo*, and there are no clinical or biological markers that provide accurate estimates for individual women.

In his 1951 discussion on the drawbacks of his methodology for estimating follicle populations by analysing slides of ovarian tissue, Block [1] provided the motivation for this paper:

> ...The distribution of these follicles in human ovaries is so uneven that reliable values can not be obtained until *all* the follicles are counted. This requires complete serial sectioning, which for a woman of fertile age means one thousand five hundred to two thousand five hundred 20μm sections per ovary. Under these circumstances any large scale investigation is impracticable....

The central claim of this paper is that advances in technical histopathology (staining and digital imaging of tissue), together with advances in data storage & retrieval, AI search, feature recognition and image analysis, make the automation of follicle counting a practical – if far from trivial – undertaking in the twenty-first century.

1.1 The importance of determining follicle populations

Human ovarian function is not well understood. The best available current models that associate primordial follicle population with age are severely flawed: they are based on archaic histopathological methods used on insufficient numbers of ovaries to obtain statistically significant results. A better model of primordial follicle with age would be important in two areas: the prediction of age at menopause, and the quantification of the effects of treatment for cancer on the fertility of survivors.

1.2 Predicting age at menopause

Wallace and Kelsey have shown a strong correlation between ovarian reserve and ovarian volume [2]. However, ovarian reserve is currently defined by the Faddy-Gosden model for follicle population with age [3], which, although taking into account all available histopathological data, is not known to be accurate. The Faddy-Gosden model for population y at age x is given by

$$\frac{dy}{dx} = -y(0.0595 + \frac{3716}{11780 + y}), \quad y(0) = 701,200$$

A population of less than 1,000 is associated with ovarian failure, i.e. menopause.

By providing a better model, using digital image analysis for cell classification and enumeration, we can improve candidate methods for determining age at menopause.

1.3 Late effects of cancer treatment

The Faddy-Gosden method has been used to quantify the radiosensitivity of the human oocyte [4], to determine the Effective Sterilising Dose for human females [5] (the dose that will immediately sterilise 97.5% of women or girls undergoing radiotherapy), and to quantify the effects of radiotherapy on a treated ovary [6]. Various cytotoxic agents are used as cancer treatments, with some (notably platinum based agents) being more toxic to ovarian cells than others. Before we can hope to quantify the damage done by such chemotherapies to the ovaries, we need a better understanding of the ovarian reserve in healthy women. If we do not have a good model for untreated females, we can not obtain decent results for survivors of cancer.

2 Background and review of current methods

2.1 The primordial follicle

Primordial follicles are considered the fundamental reproductive units of the ovary because they give rise to all dominant follicles, and therefore to all menstrual cycles. The ovarian primordial follicle (Figure 1) consists of a primary

Fig. 1. A human primordial follicle. The oocyte with its nucleus is surrounded by a single layer of granulosa cells, both of which are enclosed within a basal lamina. The diameter of a human primordial follicle is 30μm – 50μm. Reproduced with permission from the Labortorio de Anatomia Patalogica y Citologia de Adultos, Hospital de la Mujer, Centro Hospitalario Pereira Rossell, Montevideo, Uruguay.

oocyte surrounded by a layer of spindle-shaped granulosa cells demarcated by a basal lamina. An oocyte is considered primary if it has not yet undergone its first meiotic division. The human ovary contains a fixed pool of primordial follicles which declines with increasing age culminating in the menopause at an average age of 50-51 years. Primordial follicles do not have an independent blood supply and thus have limited access to the endocrine system. All primordial follicles (oocytes) are formed in the human foetus between the sixth and the ninth month of gestation. As a result, all oocytes capable of participating in reproduction during a woman's life are present in the ovaries at birth. The number of eggs or primordial follicles in a woman's ovaries constitutes her ovarian reserve.

2.2 Existing techniques

We now briefly summarise historical approaches and the state of the art in estimating ovarian reserve. A more detailed discussion of the difficulties involved in counting ovarian follicles is given in [7].

In the 1950s, Erik Block presented the first detailed methodology for counting and classifying ovarian follicles [1], applied the results to ovaries taken from 71 women and girls aged 6–44 years who had died in accidents [8], and further

Fig. 2. A cluster of primordial follicles. Reproduced with permission from the Laboratorio de Anatomia Patalogica y Citologia de Adultos, Hospital de la Mujer, Centro Hospitalario Pereira Rossell, Montevideo, Uruguay.

applied the results to ovaries from 10 infants who died at or close to birth [9]. His methods were the best available then, and have formed the basis for other studies on mammalian ovarian reserve. Block rejected the use of Wicksell's 1925 solutions of the corpuscle problem, given as the integral equation

$$1 - G_3(x) = c \int_x^\infty \frac{1}{\sqrt{t^2 - x^2}} dG_2(t), \quad x > 0$$

where G_3 is the probability distribution of an aggregate of non-overlapping balls (i.e. follicles) of random sizes scattered in a three-dimensional specimen (i.e. an ovary) to be determined from G_2, a similar probability distribution, but for discs in a plane (i.e. a sectioned ovary), and the constant c is defined as

$$c^{-1} = \int_0^\infty t^{-1} dG_2(t).$$

Although an attractive option, the method of efficiently solving such equations using numerical quadrature of Gauss-Chebyshev type had not been fully developed at that time. Instead, Block adapted a formula due to Floderus (1944) for the estimation of the total number of cells in a volume:

$$N = n. \frac{a}{a + d - 2h}$$

where N is the required number, n is an approximation calculated by counting cell nuclei appearing in histopathological slides, a is the section thickness, d is the average cell diameter, and h is the smallest height of cell tissue that could be

identified in a slide. Essentially, this method corrects for the same cell appearing on more than one slide image. Block's work was groundbreaking and, given the equipment and facilities available to him, an outstanding contribution to our knowledge of ovarian function with respect to age. However, as Block himself acknowledged, his studies suffered from significant drawbacks:

- the value for d was supposed to be the true mean follicular volume, but was in fact a representative mean taken from a small sample;
- although each ovary was sectioned at 20μm intervals, data was taken from every 200th section (with 99.5% of all available data not being analysed);
- only a small number of ovaries were studied, with not enough ovaries from women, girls or infants being available to provide suitable power and precision for statistical analysis;
- although the distribution of follicles is far from even, all estimates were based on extrapolation of distributions found from a small sample of ovarian tissue;
- the ovaries were preserved in formaldehyde post-autopsy, which damaged some follicles and shrank the volume, thus introducing serious errors into his calculations.

Despite these drawbacks, other studies have either used Block's methods, or developed similar models for estimating populations. For example an important study in 2001 took a representative sample of tissue, obtained a mean primordial follicle volume from fewer than 100 samples, and estimated the total population by assuming that the volumes and distribution were similar throughout the ovary [10]. In 2003 a new mathematical model was presented for population density [11]. The population is given by

$$n^{\gamma} = \gamma.\alpha. \sum_{i=1,\gamma+1,2\gamma+1,\ldots}^{N} n_i$$

where N, is the total number of sections, n_i is the number of follicles observed in section i, γ is a correction factor that accounts for incomplete slide analysis (i.e. $\gamma = 5$ if every fifth section is used), and α is another correction factor that accounts for counting the same follicle two or three times. The formula for α is

$$\alpha = P_2.\frac{1}{2} + P_3.\frac{1}{3}, \text{ in which } P_2 = \frac{2t-d}{t}, P_3 = \frac{d-t}{t}$$

where t is the slice thickness and d is the mean diameter of a follicle. It is easy to derive a formula for the density of follicles in a given volume from these equations.

As recently as 2006, a similar stereological method was utilised to estimate the total number of oogonia in eleven foetal ovaries [12]. These approaches are an improvement on the methodology of Block: staining and imaging has improved, and the standard errors in the formulae used are likely to be lower (although no study providing mathematical or empirical evidence for this has been published). However, several of the shortcomings of Block's method remain:

– a mean follicular volume, d, is calculated from a sample image and used throughout all subsequent calculations;
– not all slides are analysed for data – every second or every fifth slide is observed;
– slide analysis is performed **by hand** – observers click on each follicle appearing in a slide image using a mouse, then move on to the next one;
– all population counts assume a regular distribution function for follicles, even though it is well known that follicles are **not** distributed regularly.

In the following Section, we describe a new procedure that eliminates each of these problems. The procedure depends heavily on advanced computer architectures, algorithms and AI techniques: we aim to eliminate errors by careful automation of the entire process.

3 A new methodology for determining primordial follicle population

The methodology comprises three stages:

1. Histopathological preparation of ovarian sections, and digital photography;
2. Automated feature recognition software that delineates each section of primordial follicle appearing in an image;
3. Automated analysis – across images – to determine which follicles appear in more than one image, so that an accurate count is returned.

We now describe each of these areas in more detail.

3.1 Histopathology

Current techniques for obtaining (i) uniform sections of ovarian tissue, (ii) staining the tissue, and (iii) obtaining high quality digital photographs, are adequate for our purposes. There are, however, some technical details involved. Firstly, the sections should be at 20μm intervals, since primordial follicles have diameters ranging from 30μm through 50μm. This interval width ensures that each distinct primordial follicle will appear in one, two or three images, and not in four or more. It is not coincidental that this width was suggested as an ideal by Block in 1951.

Secondly, our preliminary results suggest that haematoxylin and eosin stain will work well: the haematoxylin imparts a blue colour to cell nuclei, and a pink colour to other tissue. We can then use the single layer of granulosa cells (with small blue nuclei) to identify and classify primordial follicles. Staining follicular nuclei blue is also advantageous when follicles are arranged in such a way that two distinct small follicles appear to represent one large one: by counting nuclei we can differentiate these cases. However, it should be noted that this apparent benefit is clear when using human eyes to identify follicles in images. There may

be other stains (for example, cajal gallego or azan mallory) that are more suited to automated feature recognition.

The third technical aspect is compression. Image quality is primarily dependent on the compression technique and level of compression, and also directly affects the file size. If an image is saved in TIFF, the quality is the highest and as there is no compression that affects the image quality, the file size is also very big. Normal histological practice for a study involving 1,500 – 2,500 images **per ovary**, would be to compress to JPEG format to reduce data storage requirements. For our method the highest possible amount of data per image is needed, and hence no compressed formats.

A fourth consideration is the potential use of immunofluorescence techniques. Primordial follicles contain lower levels of follicle stimulating hormone (FSH) than other follicles, and there are other endocrine markers (anti-M ullerine hormone, estradiol and inhibin B) that could be used to classify primordial follicles. This would involve the development of antigens that, when used to stain ovarian tissue, would make the primordial follicles stand out under certain spectral conditions. This, however, is beyond the scope of this paper.

3.2 Image Analysis

Suppose that we have a set of images of a sectioned ovary. Figure 2 is an example of a fragment of such an image. The next stage is to automatically identify, select and count primordial follicles. The steps involved in this object-oriented image analysis process are (i) obtain a segmented image, (ii) convert to a classified image, (iii) obtain feature statistics.

For tissue segmentation, image objects are classified into histopathology classes from spectral and shape characteristics and spatial relationships between tissue objects. Primordial follicles are defined as an oocyte surrounded by a single layer of flat granulosa cells, and are know to vary in size from 30μ m tp 50μm. There are no ovarian cells that match this description: granulosa cells are much smaller, and primary and secondary follicles are larger (more than 70μm) and have surrounding layers of cuboid or elongated ganulosa cells. The only possible confounders are blood vessels sections, but these clearly have no nuclei, and therefore can be automatically rejected. Hence we have a clear set of morphometric features that define primordial follicles, and only primordial follicles. We can therefore apply machine-learning techniques to get accurate classification from segmented images.

We use a modified version of a training algorithm that maximizes the margin between the training patterns and the decision boundary [13]. The algorithm proceeds as follows:

- the feature with the lowest contribution to the model is dropped and a new model is constructed on the remaining features;
- this procedure is repeated until there are no more features left for an algorithm model to be trained on;

– at this point, the model with the highest fitness is selected. (In the case of multiple models with equal values of the fitness, the model with the fewest features is selected.)

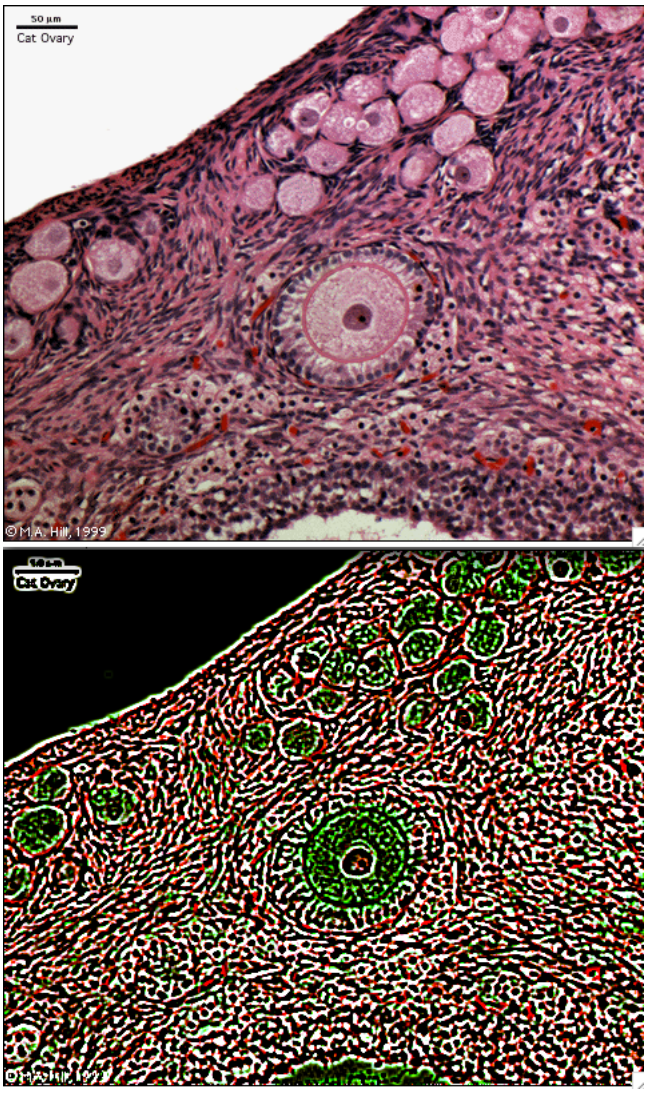

Fig. 3. Above: an original image of feline ovarian tissue stained with haematoxylin and eosin. The large follicle in the centre is a primary follicle; the smaller follicles are primordial. Image reproduced with permission from McGraw-Hill. Below: the same image after (i) smoothing, (ii) convolution, and (iii) RG2B colocalisation. The follicles are clearly identified from background tissue.

Using trained classifiers to select the primordial follicles appearing in each image (together with a binary indicator of the absence or otherwise of a nucleus in each primordial follicle), we can now decide whether a given follicle appears in one, two or three consecutive images.

3.3 Counting Follicles

We model the problem as a constraint satisfaction problem (CSP). These are generalisations of Propositional Satisfiability (SAT) problems, and the class of CSPs is therefore NP-complete. Efficient heuristics exist for these problems, in general involving backtrack search for solutions, with local consistency propogation algorithms used to reduce the sizes of allowed domains, hence pruning search. CSPs and techniques for solving them are described fully in [14].

Definition 1. *A CSP P is a triple $(\Delta, \mathcal{D}, \mathcal{C})$, where Δ is a finite indexed set of variables $x_1, x_2, \ldots, x_n$, each of which has finite domain of possible values $D_i := \mathrm{Dom}(x_i) \subseteq \Lambda$. The set $\mathcal{D} = \{D_i : 1 \leq i \leq n\}$, and the set $\mathcal{C}$ is a finite set of constraints on the variables in Δ. A* solution *to a CSP is an instantiation of all of the variables in Δ such that all of the constraints in $\mathcal{C}$ are satisfied.*

Suppose that we have a region R_k of pixels in image k that has been classified as a section of a primordial follicle. We define the set of variables $\Delta = \{R_{k-2}, R_{k-1}, R_k, R_{k+1}, Rk+2\}$, each having domain $\mathcal{D} = \{-1, 0, 1\}$. Domain value 0 represents the variable not being classified as containing a primordial follicle. Domain value 1 represents a positive classification **with** the nucleus of the oocyte appearing in the image section. Domain value -1 represents a positive classification **without** the nucleus appearing. We have a single constraint, $\mathcal{C}$ which is the disjunction of the following clauses:

$$R_k = 1 \wedge R_{k-1} = 0 \wedge R_{k+1} = 0 \tag{1}$$
$$R_k = 1 \wedge R_{k-1} = -1 \wedge R_{k-2} = 0 \wedge R_{k+1} = 0 \tag{2}$$
$$R_k = 1 \wedge R_{k-1} = -1 \wedge R_{k-2} = -1 \wedge R_{k+1} = 0 \tag{3}$$
$$R_k = 1 \wedge R_{k-1} = 0 \wedge R_{k+1} = -1 \wedge R_{k+2} = 0 \tag{4}$$
$$R_k = 1 \wedge R_{k-1} = 0 \wedge R_{k+2} = -1 \wedge R_{k+1} = -1 \tag{5}$$
$$R_k = 1 \wedge R_{k-1} = -1 \wedge R_{k-2} = 0 \wedge R_{k+1} = -1 \wedge R_{k+2} = 0 \tag{6}$$
$$R_k = -1 \wedge R_{k-1} = 1 \wedge R_{k-2} = 0 \wedge R_{k+1} = -1 \wedge R_{k+2} = 0 \tag{7}$$
$$R_k = -1 \wedge R_{k-1} = -1 \wedge R_{k-2} = 0 \wedge R_{k+1} = 1 \wedge R_{k+2} = 0 \tag{8}$$
$$R_k = -1 \wedge R_{k-1} = 0 \wedge R_{k+1} = -1 \wedge R_{k+2} = 1 \tag{9}$$
$$R_k = -1 \wedge R_{k-1} = 0 \wedge R_{k+1} = 1 \wedge R_{k+2} = -1 \tag{10}$$
$$R_k = -1 \wedge R_{k-1} = 0 \wedge R_{k+1} = -1 \wedge R_{k+2} = 0 \tag{11}$$
$$R_k = -1 \wedge R_{k+1} = 0 \wedge R_{k-1} = -1 \wedge R_{k-2} = 1 \tag{12}$$
$$R_k = -1 \wedge R_{k+1} = 0 \wedge R_{k-1} = 1 \wedge R_{k-2} = -1 \tag{13}$$
$$R_k = -1 \wedge R_{k+1} = 0 \wedge R_{k-1} = 1 \wedge R_{k-2} = 0 \tag{14}$$

Clauses 1 through 6 represent an oocyte nucleus in R_k, with the associated primordial follicle appearing in zero, one or two further consecutive images. Clauses 7 through 14 represent a single nucleus, not in R_k, in a primordial follicle involving R_k. The constraint enforces the basic morphological requirements that each primordial follicle must satisfy: in can appear in at most three consecutive images, and must contain exactly one nucleus.

A solution to a problem of this kind allows us to label the image sections that identify a unique primordial follicle, and therefore count each distinct follicle.

Strictly speaking, we do not need to formulate these problems as CSPs, since we only need to check for satisfaction in one of fourteen distinct cases for each R_k tested. The reason for using CSPs is that non-solutions and multiple solutions can be used to identify and deal with possible false negatives and false positives. These can arise when primordial follicles are clustered together, with more than three consecutive similar regions being classified as containing follicular tissue. Hence, this method is predicted to have high sensitivity **and** high specificity. False negatives are unlikely because cell slices of the correct size containing a nucleus are easily identifiable. It may be that any non-nucleus containing sections are mis-classified, but this will not affect the overall count, only whether the nucleus slice is associated with other slices. False positives are also unlikely: the only ovarian cells in the given size range are primordial follicles.

4 Implementation and evaluation

We now describe preliminary prospective data involving the image analysis and follicle counting stages detailed in the previous section. For training purposes we use images obtained from other studies, most importantly from a study involving the cryopreservation of ovarian tissue prior to treatment for cancer [15].

4.1 Image analysis

Our implementation involves use of ImageJ, a microscopy image analysis toolkit developed by the National Institute of Health, USA [16]. The procedure for an image of ovarian tissue is:

1. digitally smooth the image;
2. perform convolution on the result, giving green and blue follicular tissue and red and blue background tissue;
3. perform RGB colocalisation, removing blue pixels strongly associated with red or green pixels, giving green follicular tissue and red/pink background tissue;
4. convert to 8-bit grayscale, and obtain watershed segmentation so that each follicle section is defined as a distinct region;
5. remove all (primary, secondary and antral) follicles greater than 50μm in largest diameter.

A similar process associates nuclei with primordial follicles as required. Figure 3 shows the effect of stages 1–3. Individual follicles are clearly defined, with nuclei appearing as red centres of green oocytes. Remarkably little training has been needed so far, with primordial follicles being clearly identified from the outset. Using advanced hardware we can process and store all information from an average human ovary (2,000 sections) in less than one day.

4.2 Follicle counting

The CSPs associated with each region of an image associated with a primordial follicle are easy to solve, but there are many of them. We have implemented a solver in the prolog-based Constraint Programming toolkit Eclipse [17], using symmetry-breaking techniques developed in [18]. Using random data (i.e. sequences of regions that are not taken from images of tissue, but derived randomly using parameters obtained from our initial results from the image analysis component of the methodology) we can solve each CSP quickly (several microseconds on average), using the results to obtain accurate counts. A major overhead is the low-level manipulation of images to ensure that each similar region from 5 consecutive images is available for input to the CSP. For a study involving 1,500 – 2,500 images per ovary it appear that distributing this manipulation across multiple computers – using the Mac OS XGrid controller to split jobs into tasks, resubmit failures and retrieve results – is likely to be beneficial.

5 Conclusions and Further Work

We have presented a novel technique for classifying and counting primordial follicles. The technique represents an improvement on existing methodologies. It does not rely on estimates of population distributions based on sample data, nor does it rely on follicle classification by hand, but instead performs advanced digital analysis of each image of every section of an ovary. Using this technique we can obtain a better understanding of human ovarian function, leading to more accurate prediction of age at menopause, and improved quantification of the effects of radio- and chemotherapy of the fertility of cancer survivors. Preliminary implementation and empirical evaluation demonstrates the basic efficacy of the technique, which involves a sophisticated combination of machine learning, intelligent feature recognition, and AI search.

Future work involves the deployment of the method using data obtained from the histopathology of ovaries obtained from planned oophorectomies. It remains to be determined which choice of stain, magnification and image resolution will most suitable for automated image analysis and follicle counting. Moreover, the possibility of staining with specific antigens that transform primordial follicle sections (and only primordial follicle sections) into an easily detectable colour when photographed, remains unexplored.

References

1. E. Block. Quantitative morphological investigations of the follicular system in women; methods of quantitative determinations. *Acta Anatomica*, 12(3):267–85, 1951.
2. W.H.B Wallace and T.W. Kelsey. Ovarian reserve and reproductive age may be determined from measurement of ovarian volume by transvaginal sonography. *Human Reproduction*, 19(7):1612–1617, 2004.
3. M.J. Faddy and R.G. Gosden. A model conforming the decline in follicle numbers to the age of menopause in women. *Human Reproduction*, 11(7):1484–86, 1996.
4. W.H.B Wallace, A.B. Thomson, and T.W. Kelsey. The radiosensitivity of the human oocyte. *Human Reproduction*, 18(1):117–121, 2003.
5. W.H.B Wallace, A. Thomson, F. Saran, and T.W. Kelsey. Predicting age of ovarian failure after radiation to a field that includes the ovaries. *International Journal of Radiation Oncology*Biology*Physics*, 62(3):738–744, 2005.
6. T.W. Kelsey. Efectos de la radioterapia sobre el ovario y el útero. *La opinión de los expertos – Actualidades y controversias en ginecologóa*, 36(8):20–23, March 2007.
7. J.L. Tilly. Ovarian follicle counts–not as simple as 1, 2, 3. *Reproductive Biology and Endocrinology*, 1(11), 2003.
8. E. Block. Quantitative morphological investigations of the follicular system in women; variations at different ages. *Acta Anatomica*, 14(1):108–23, 1952.
9. E. Block. A quantitative morphological investigation of the follicular system in newborn female infants. *Acta Anatomica*, 17(3):201–6, 1953.
10. M.E. Pepling and A.C. Spradling. Mouse ovarian germ cell cysts undergo programmed breakdown to form primordial follicles. *Developmental Biology*, 234(2):339–51, 2001.
11. K.L.T. Schmidt, A.G. Byskov, A. Nyboe Andersen, J. Mller, and C. Yding Andersen. Density and distribution of primordial follicles in single pieces of cortex from 21 patients and in individual pieces of cortex from three entire human ovaries. *Human Reproduction*, 18(6):1158–1164, 2003.
12. E. Bendsen, A.G. Byskov, C. Yding Andersen, and L.G. Westergaard. Number of germ cells and somatic cells in human fetal ovaries during the first weeks after sex differentiation. *Human Reproduction*, 21(1):30–35, 2006.
13. B. E. Boser, I. Guyon, and V. Vapnik. A training algorithm for optimal margin classifiers. In *Computational Learing Theory*, pages 144–152, 1992.
14. F Rossi, P van Beek, and T Walsh, editors. *The Handbook of Constraint Programming*. Elsevier, 2006.
15. F. Cóppola, J. Martínez Torena, A. Domínguez, S. Parada, H. Pérez Campos, Y.Alfonzo, L. De Dios, and R. Aguirre. Reestablecimiento de la función endócrina de implantes ováricos heterotrópicos no pediculizados. *Revista Médica del Uruguay*, 23(1):56–63, 2007.
16. W.S. Rasband. *ImageJ*. U.S. National Institutes of Health, Bethesda, Maryland, 1997. `(http://rsb.info.nih.gov/ij/)`.
17. M. G. Wallace, S. Novello, and J. Schimpf. ECLiPSe : A platform for constraint logic programming. *ICL Systems Journal*, 12(1):159–200, May 1997.
18. T.W. Kelsey, S.A. Linton, and C.M. Roney-Dougal. New developments in symmetry breaking in search using computational group theory. In B. Buchberger and J. A. Campbell, editors, *AISC*, volume 3249 of *Lecture Notes in Computer Science*, pages 199–210. Springer, 2004.