

2016

Probabilistic Analysis of Polyovulation Frequencies

Tiffany N. Kolba

Valparaiso University, tiffany.kolba@valpo.edu

Kaylyn Banaszak

Valparaiso University, Kaylyn.Banaszak@valpo.edu

Anna Kaniewski

Valparaiso University, anna.kaniewski@valpo.edu

Follow this and additional works at: http://scholar.valpo.edu/math_stat_fac_pubs

Recommended Citation

Kolba, T.N., Banaszak, K., and Kaniewski, A. (2016). Probabilistic analysis of polyovulation frequencies. *Spora: A Journal of Biomathematics*, 2(1). 18-26.

This Article is brought to you for free and open access by the Department of Mathematics and Statistics at ValpoScholar. It has been accepted for inclusion in Mathematics and Statistics Faculty Publications by an authorized administrator of ValpoScholar. For more information, please contact a ValpoScholar staff member at scholar@valpo.edu.

2016

Probabilistic Analysis of Polyovulation Frequencies

Tiffany N. Kolba

Valparaiso University, tiffany.kolba@valpo.edu

Kaylyn Banaszak

Valparaiso University, Kaylyn.Banaszak@valpo.edu

Anna Kaniewski

Valparaiso University, anna.kaniewski@valpo.edu

Follow this and additional works at: <http://ir.library.illinoisstate.edu/spora>

Recommended Citation

Kolba, Tiffany N.; Banaszak, Kaylyn; and Kaniewski, Anna (2016) "Probabilistic Analysis of Polyovulation Frequencies," *Spora: A Journal of Biomathematics*: Vol. 2: Iss. 1, Article 3.

Available at: <http://ir.library.illinoisstate.edu/spora/vol2/iss1/3>

This Research is brought to you for free and open access by ISU ReD: Research and eData. It has been accepted for inclusion in Spora: A Journal of Biomathematics by an authorized administrator of ISU ReD: Research and eData. For more information, please contact ISURed@ilstu.edu.

Probabilistic Analysis of Polyovulation Frequencies

Tiffany N. Kolba^{1,*}, Kaylyn Banaszak¹, Anna Kaniewski¹

*Correspondence:
Prof. Tiffany Kolba, Dept. of
Mathematics and Statistics,
Valparaiso University,
1900 Chapel Drive,
Valparaiso, IN 46383-6493,
USA
tiffany.kolba@valpo.edu

Abstract

Polyovulation is the production of more than one ovum, or egg, during a single menstrual cycle. This paper examines the probability of the human ovarian system ovulating k eggs during a single cycle, for $k \geq 0$. In order to obtain precise estimates for the probability of polyovulation, we use U.S. birth data from the 1950s (before the introduction of artificial reproductive technologies). However, to utilize birth data, we model the various processes that eggs undergo in order to result in a live birth, including fertilization, possible division, implantation, and potential miscarriage. Our model produces novel estimates for the probability that a fertilized egg divides, as well as for the zygosity type frequencies of twins, triplets, and quadruplets.

Keywords: probability, polyovulation, zygosity, twins, birth data

1 Introduction

Probabilistic techniques have been applied to the study of multiple births at least since 1901 when Weinberg developed the Weinberg Differential Rule to estimate the population relative frequencies of identical, i.e. monozygotic, versus fraternal, i.e. dizygotic, twins [10]. Identical twins possess the same DNA, while fraternal twins do not, and hence proper classification can have important health repercussions. Up through the early 1900s, the prevailing medical practice was to classify twins of the same sex as identical if the twins shared a placenta and as fraternal if each twin had its own placenta. However, the Weinberg Differential Rule indicated that classification based upon number of placentas resulted in an overabundance of same sex fraternal twins compared to what would be expected from basic probability. This realization eventually led the medical community to conclude that identical twins could indeed have separate placentas and increased proper classification of twins.

To illustrate how the Weinberg Differential Rule works, we consider U.S birth data from 1952–1954. Due to constraints of available data, only Caucasian births are used in the analysis. Birth records indicate that there were 32,923 sets of opposite sex twins and 72,547 sets of same sex twins [1]. All of the opposite sex twins must be fraternal since identical twins share the same DNA and hence must be the same sex. Assuming that the sex of each twin is independent and that males and females are equally likely, assumptions that are fairly reasonable, 50% of fraternal twins should be the same sex and 50% should be

opposite sex. Hence, if there were 32,923 sets of opposite sex fraternal twins, there should be approximately 32,923 sets of same sex fraternal twins. This gives a total of 65,846 sets of fraternal twins and the remaining sets as identical. Dividing the estimates by 105,470, the total number of sets of twins, gives an estimate that 62.4% of twins are fraternal and 37.6% of twins are identical. Now for a particular set of same sex twins, the only way to know for sure whether they are identical or fraternal is to perform a DNA test, but the Weinberg Differential Rule gives a simple, yet powerful, method for estimating population frequencies.

Since the Weinberg Differential Rule was developed in 1901, there has been extensive research into the phenomenon of twins and higher order multiples [3, 5], but limited research into the phenomenon of polyovulation, which is one of the biological mechanisms that produces multiple births. In particular, polyovulation, also known as superovulation, is the production of more than one ovum, or egg, during a single menstrual cycle. This paper focuses on estimating the probability of the human ovarian system ovulating k eggs in a single cycle for $k \geq 0$.

In order to directly measure the number and locations of eggs ovulated, ultrasound images of the ovaries can be utilized to count the number of corpora lutea. The corpus luteum is the remnant of the follicle where the egg was released from the ovary and is about 1–3 cm, and hence visible to an ultrasound [9]. However, the problem with using the number of corpora lutea to estimate probabilities of polyovulation is that polyovulation is a fairly rare event and there simply is not adequate ultrasound data available. Hence, we instead use birth data in order to estimate polyovulation frequencies. In particu-

¹Department of Mathematics and Statistics, Valparaiso University, Valparaiso, IN

Birth Type	Raw Number	Relative Frequency
Singletons	10,283,346	0.989754
Twin Sets	105,470	0.010151
Triplet Sets	973	0.000094
Quadruplet Sets	11	0.000001

Table 1: U.S. Birth Data from 1952–1954

lar, we use the U.S. Caucasian birth data from 1952–1954 that was mentioned previously. We utilize data from the 1950s since that is before the introduction of artificial reproductive technologies, such as ovulation inducing drugs and in-vitro fertilization, and we are primarily interested in the phenomenon of *spontaneous* polyovulation. Table 1 lists the raw numbers and relative frequencies of singletons, twin sets, triplet sets, and quadruplet sets in the data set. No quintuplets or higher order multiples were recorded [1].

Now, we cannot simply use the birth relative frequencies as our estimates of polyovulation frequencies because there are many biological processes that occur in between ovulation and birth. In particular, once the eggs are ovulated, they must be fertilized in order to become zygotes. Then the zygotes have to implant successfully in the uterus of the mother, and can possibly divide before or after doing so. The implanted embryos then have to survive the possibility of a potential miscarriage in order to actually result in live births. In the following section, we will model each of these stages in succession. In the course of our modeling, we produce novel estimates for the probability that a zygote divides, as well as for the zygosity type relative frequencies (i.e. proportion of identical versus fraternal) of twins, triplets, and quadruplets, which are given in Section 3 along with the final results for the polyovulation probabilities.

2 From Ovulation to Birth

In this section we model the four stages of ovulation, fertilization, division and implantation, and live birth in order to eventually utilize the birth type relative frequencies to achieve our goal of estimating the polyovulation frequencies.

2.1 Ovulation Model

Let O_k denote the event that exactly k ova are ovulated in one menstrual cycle for $k \geq 0$. It is the probabilities of these events that is our main goal to estimate. Now for simplicity we will assume that ovulation of 5 or more eggs in a single cycle has probability zero, i.e. $P(O_k) = 0$ for $k \geq 5$. While these probabilities are certainly not exactly zero, our birth data set had no observed quintuplets or

higher, so it is not feasible for us to estimate the probabilities of these rare events. Since we are using birth data for our estimates, these estimates are conditional on the event that there was at least one egg ovulated. Hence, we will need a separate source for the estimate of zero eggs ovulated, called anovulation. The literature estimates the probability of anovulation in fertile women to be around 0.007 [6]. Now since $P(O_0) = 0.007$ and

$$P(O_0) + P(O_1) + P(O_2) + P(O_3) + P(O_4) = 1,$$

we only need three additional linearly independent equations in order to solve for all the unknown polyovulation probabilities.

2.2 Fertilization Model

Let Z_k denote the event that exactly k zygotes, or fertilized eggs, are produced in one cycle for $k \geq 0$. We will assume that each egg is fertilized independently of all other eggs with probability p_f . The literature estimates the probability of fertilization to be around 0.3 [7, 9]. The assumption of independence is supported by the work of Tong et al. [9], which analyzed ultrasound data of pregnant women exhibiting double ovulation. Because of the independence assumption, we can model the Z_k 's in terms of the O_k 's and p_f using the following equations:

$$\begin{aligned} P(Z_0) &= P(O_0) + P(O_1)(1 - p_f) + P(O_2)(1 - p_f)^2 \\ &\quad + P(O_3)(1 - p_f)^3 + P(O_4)(1 - p_f)^4 \\ P(Z_1) &= P(O_1)p_f + 2P(O_2)p_f(1 - p_f) \\ &\quad + 3P(O_3)p_f(1 - p_f)^2 + 4P(O_4)p_f(1 - p_f)^3 \\ P(Z_2) &= P(O_2)p_f^2 + 3P(O_3)p_f^2(1 - p_f) \\ &\quad + 6P(O_4)p_f^2(1 - p_f)^2 \\ P(Z_3) &= P(O_3)p_f^3 + 4P(O_4)p_f^3(1 - p_f) \\ P(Z_4) &= P(O_4)p_f^4. \end{aligned}$$

Note that the coefficient of each term in the above equations is a combination of the form $\binom{n}{k}$ where we are counting the number of ways to choose the k eggs that are fertilized from the n eggs that were ovulated. In the following subsections we will work towards finding three linearly independent equations involving the Z_k 's, which can then be combined with the previous equations in order to solve for the polyovulation probabilities.

2.3 Division and Implantation Model

We now model the process by which zygotes possibly divide and implant in order to reach the embryonic stage. Figure 1 illustrates the possible configurations in which a certain number of zygotes can result in a certain number of embryos through division. As mentioned previously for twins, there is the case of identical twins, also called monozygotic twins, which result from the division of a single zygote, and the case of fraternal twins, also called dizygotic twins, which result from two separate zygotes with no divisions. In general, the zygosity of a set of multiples is defined to be the number of zygotes from which they derived. Hence, triplets can be either monozygotic, in which all three are identical, dizygotic, in which case there is one identical pair, or trizygotic, in which all three are fraternal. Likewise, quadruplets can be monozygotic, dizygotic, trizygotic, or quadrazygotic. In the case of monozygotic quadruplets, there are two different possible division sequences. There could be a primary division, followed by a secondary division of one of the zygotes, followed by a tertiary division. Or there could be a primary division followed by two secondary divisions. Modeling these two cases separately is important because they have different coefficients for how many similar configurations exist with equal probability by symmetry, which are indicated in Figure 1 by the boxed factors. Similarly, dizygotic quadruplets could correspond to quadruplets consisting of two identical pairs, or a set of three identical embryos with a fourth non-identical.

The timing of division is also important to model, as that will affect both the implantation probabilities and the miscarriage probabilities. An “early” division in days 1–3 after fertilization will result in embryos with separate placentas and separate amniotic sacs. A “middle” division in days 4–8 after fertilization will result in a shared placenta, but separate amniotic sacs. Implantation occurs around day 8 after fertilization, and a “late” division after that (up until around day 14 after fertilization) will result in a shared placenta and a shared amniotic sac, which has a very high risk of miscarriage [5]. The chorionicity of a set of multiples is defined to be the number of distinct placentas they possess, and the amnionicity of a set of multiples is defined to be the number of distinct amniotic sacs. Since embryos with separate placentas cannot share an amniotic sac and embryos deriving from separate zygotes cannot share a placenta, the amnionicity of a set of multiples is always greater than or equal to the chorionicity, which is always greater than or equal to the zygosity. Figure 1 illustrates all the possible zygosity and chorionicity combinations for each number of embryos, with the placentas indicated by the larger circles. The different amnionicity possibilities are not illustrated due to the large number of combinations, but

any embryos sharing a placenta could share an amniotic sac or have distinct amniotic sacs, based upon whether there was a middle or late division.

For the specific assumptions of our division and implantation model, we will assume that each zygote divides independently of all other zygotes. We also assume that once a zygote divides, it resets and has the same probability of subsequent divisions as a zygote that has not yet divided. We will assume that divisions that result in an overall total of more than four offspring at any given time have a negligible probability and hence will be set equal to 0. As for implantation, it is the placenta that implants in the uterus of the mother [7], so we model the implantation of the placentas, rather than of the individual zygotes. We assume that each placenta implants independently of all other placentas with probability p_i , which biological evidence suggests is reasonable [9]. Now the last assumption is that a shared placenta has the same implantation probability as an individual placenta. The validity of this assumption is not addressed in the biological literature and has a peculiar probabilistic result. Because an early division results in two separate placentas, the probability that both embryos implant will be p_i^2 , whereas since a middle division results in a shared placenta, the probability that both embryos implant is p_i , which is much greater than p_i^2 . Hence, it seems biologically plausible that a shared placenta might have a different implantation probability than an individual placenta, and our hope is that this work will perhaps motivate further research in the biological community about implantation rates for individual versus shared placentas. Regardless, the estimates from literature for the probability of implantation are around 70% [7] and, thus, we set $p_i = 0.7$.

Let δ_e denote the probability of an early division, resulting in separate placentas and separate amniotic sacs, δ_m denote the probability of a middle division, resulting in a shared placenta, but separate amniotic sacs, and δ_l denote the probability of a late division, resulting in a shared placenta and shared amniotic sac. The probability that a zygote does not divide is $(1 - \delta_e)(1 - \delta_m)(1 - \delta_l)$. No direct estimates for these division probabilities exist in the literature. Rather, we use a novel approach to estimate the values from the Weinberg Differential Rule, as described in the Introduction, and the known proportions of identical twins who are dichorionic/diamniotic, monochorionic/diamniotic, and monochorionic/monoamniotic. These division probability estimates will be described in more detail in Section 3.1 since the values depend upon the miscarriage probabilities described in Section 2.4.

Let $E_{k,z,c,a}$ denote the event that a pregnancy consists of k embryos, with zygosity z , chorionicity c , and amnionicity a . We consider these events for $1 \leq k \leq 4$, with the restriction that $1 \leq z \leq c \leq a \leq k$, resulting in 35 pos-

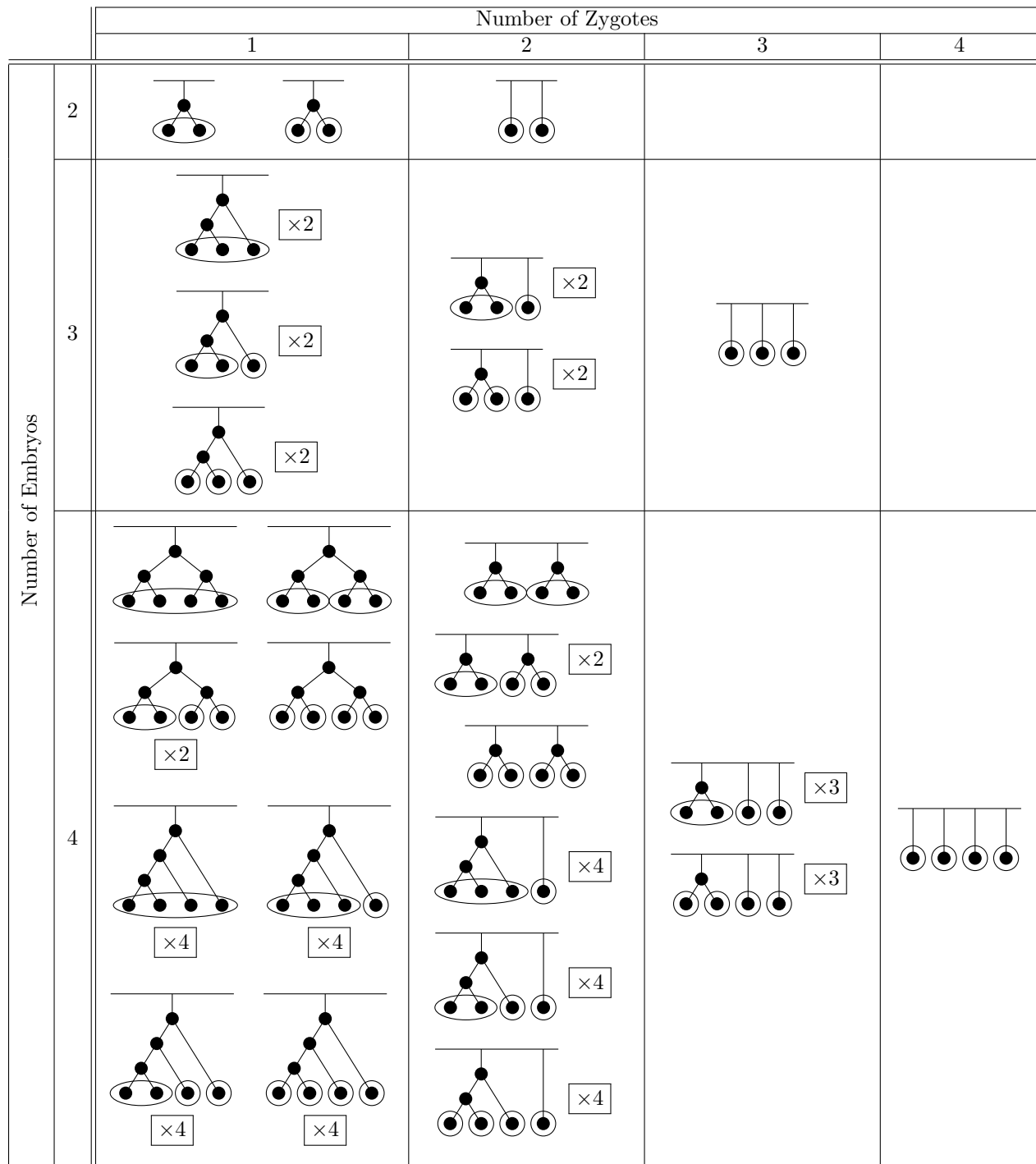


Figure 1: Possible division configurations by number of zygotes and number of embryos, with number of placentas indicated by larger circles and number of similar configurations with equal probability indicated by boxed factor.

sible events. Using the assumptions listed above regarding division and implantation, we can write $P(E_{k,z,c,a})$ in terms of the $P(Z_k)$'s, p_i , δ_e , δ_m , and δ_l . For example, the probabilities for all of the cases of quadrazygotic or trizygotic quadruplets are

$$\begin{aligned} P(E_{4,4,4,4}) &= P(Z_4)(1 - \delta_e)^4(1 - \delta_m)^4(1 - \delta_l)^4 p_i^4 \\ P(E_{4,3,4,4}) &= 3P(Z_3)\delta_e(1 - \delta_e)^4(1 - \delta_m)^4(1 - \delta_l)^4 p_i^4 \\ P(E_{4,3,3,4}) &= 3P(Z_3)(1 - \delta_e)^3\delta_m(1 - \delta_m)^4(1 - \delta_l)^4 p_i^3 \\ P(E_{4,3,3,3}) &= 3P(Z_3)(1 - \delta_e)^3(1 - \delta_m)^3\delta_l(1 - \delta_l)^4 p_i^3. \end{aligned}$$

The three cases of trizygotic quadruplets are distinguished by whether there was an early, middle, or late division; all have a coefficient of 3 because there are three choices for which of the zygotes undergoes the division. Note that in each case the power of p_i is equal to the chorionicity since p_i is the probability of implantation for a distinct placenta, rather than for a distinct embryo. The early division results in four distinct placentas and hence a p_i^4 factor, whereas the middle and late divisions result in three distinct placentas, with one of them shared by two embryos, and hence a p_i^3 factor.

As the zygosity decreases or the number of embryos decreases, there are more cases to consider. For example, the probability of trizygotic triplets is

$$\begin{aligned} P(E_{3,3,3,3}) &= P(Z_3)(1 - \delta_e)^3(1 - \delta_m)^3(1 - \delta_l)^3 p_i^3 \\ &+ 6P(Z_3)\delta_e(1 - \delta_e)^4(1 - \delta_m)^4(1 - \delta_l)^3 p_i^3(1 - p_i) \\ &+ 4P(Z_4)(1 - \delta_e)^4(1 - \delta_m)^4(1 - \delta_l)^3 p_i^3(1 - p_i). \end{aligned}$$

Note that the first term in $P(E_{3,3,3,3})$ corresponds to the case of three zygotes, none of which divide and all of which successfully implant, whereas the second term corresponds to the case of three zygotes, one of which divides early, but then one of the zygotes resulting from the division does not successfully implant. The coefficient of the second term is 6 because there are three choices for which zygote divides and two choices for which of the resulting zygotes does not implant. The third term in $P(E_{3,3,3,3})$ corresponds to the case of four zygotes where none divide, but one does not successfully implant. Although this case did originally derive from four zygotes, the triplets would still be referred to as trizygotic because the remaining offspring only derived from three zygotes. The coefficient of 4 is due to choosing which of the four zygotes does not implant. The expressions for the 30 remaining $P(E_{k,z,c,a})$ are omitted for brevity, but follow from similar logic.

2.4 Birth Model

The final stage of our model from ovulation to birth is to model the process in which a certain number of embryos result in a certain number of live births, with the possibility of miscarriage in between. Let B_k denote the event

that a pregnancy results in k live births, for $k \geq 0$. Due to our assumption that a total of more than four offspring has a negligible probability,

$$P(B_0) = 1 - P(B_1) - P(B_2) - P(B_3) - P(B_4).$$

Now, for $k \geq 1$,

$$P(B_k) = \sum_{z=1}^k \sum_{c=z}^k \sum_{a=c}^k P(B_{k,z,c,a})$$

where $B_{k,z,c,a}$ is the event that a pregnancy results in k live births with the offspring possessing zygotity z , chorionicity c , and amnionicity a . We consider $1 \leq k \leq 4$ and $1 \leq z \leq c \leq a \leq k$ for a total of 35 possibilities. We will express $P(B_{k,z,c,a})$ in terms of $P(E_{k,z,c,a})$ and miscarriage probabilities.

Let β_k denote the probability that an embryo with its own amniotic sac and a placenta supplying a total of k embryos survives to live birth and p_a denote the probability that an embryo which shares an amniotic sac with at least one other embryo survives to live birth, relative to the probability for an embryo that has its own amniotic sac. We will assume that each embryo is miscarried independently of all other embryos, which is certainly the most unrealistic of all our assumptions, but assumed for feasibility. We also assume that the probability of miscarriage does not depend upon the zygotity. This implies that identical twins with separate placentas have an equal probability of miscarriage to fraternal twins, which always have separate placentas. The biological literature largely supports this, but there is perhaps some evidence that even identical twins with separate placentas have a higher rate of miscarriage due to chromosomal problems resulting from the division of a single zygote [5].

While it is difficult to accurately estimate the survival probabilities since many miscarriages go unreported, we base our estimates for the β_k 's and p_a from values given in the work of Allen [1], which utilized the same U.S. birth data set from 1952–1954 that we use in this paper. In particular, we estimate $\beta_1 \approx 0.8571$, $\beta_2 \approx 0.7840$, $\beta_3 \approx 0.7696$, $\beta_4 \approx 0.7544$, and $p_a \approx 0.5$. Note that since we are using birth data from the 1950s, these survival probabilities are estimates for pregnancies in the 1950s; survival probabilities today are likely much higher due to advances in prenatal care.

Using our assumptions listed above, the probability of the birth of trizygotic triplets is

$$\begin{aligned} P(B_{3,3,3,3}) &= P(E_{3,3,3,3})\beta_1^3 \\ &+ P(E_{4,4,4,4})4\beta_1^3(1 - \beta_1) \\ &+ P(E_{4,3,4,4})2\beta_1^3(1 - \beta_1) \\ &+ P(E_{4,3,3,4})2\beta_1^2\beta_2(1 - \beta_2) \\ &+ P(E_{4,3,3,3})2\beta_1^2\beta_2p_a(1 - \beta_2p_a). \end{aligned}$$

The first term in $P(B_{3,3,3,3})$ comes from the case of three embryos that were trizygotic and all three survive to live birth. The second term is the case of four embryos that were quadrazygotic, but only three survive to live birth. The coefficient of 4 comes from which of the four embryos was miscarried. The last three terms come from cases of four embryos that were trizygotic. There is now only a coefficient of 2 because in order for the end result to be trizygotic, one of the two identical embryos that resulted from the division of a zygote must be miscarried. The differences between the three terms come from whether or not the two identical embryos shared a placenta or amniotic sac. The other possibilities for $P(B_{k,z,c,a})$ follow from similar logic.

Now based upon the birth type relative frequencies from the U.S. 1952–1954 birth data set given in Table 1,

$$\begin{aligned} P(B_2|B_1 \cup B_2 \cup B_3 \cup B_4) &= 0.010151 \\ &= \frac{P(B_2)}{P(B_1) + P(B_2) + P(B_3) + P(B_4)} \\ P(B_3|B_1 \cup B_2 \cup B_3 \cup B_4) &= 0.000094 \\ &= \frac{P(B_3)}{P(B_1) + P(B_2) + P(B_3) + P(B_4)} \\ P(B_4|B_1 \cup B_2 \cup B_3 \cup B_4) &= 0.000001 \\ &= \frac{P(B_4)}{P(B_1) + P(B_2) + P(B_3) + P(B_4)} \end{aligned}$$

The equations listed above produce three linearly independent equations for $P(B_1)$, $P(B_2)$, $P(B_3)$, and $P(B_4)$. We can trace back the expressions for the $P(B_k)$'s in terms of the $P(O_k)$'s working backwards from our birth model, to our division and implantation model, to our fertilization model. We can then combine these three linearly independent equations with the two linearly independent equations given in the ovulation model in Section 2.1. These five linearly independent equations can then be used to solve for the five unknown polyovulation probabilities, $P(O_0)$, $P(O_1)$, $P(O_2)$, $P(O_3)$, and $P(O_4)$, using basic linear algebra techniques. The final estimates are given in Section 3.2.

3 Probability Estimate Results

In this section we present our key results for the estimates of the division probabilities and polyovulation probabilities. In addition, because our model described in the previous section tracked the zygoty, chorionicity, and amnionicity types of all multiple births, we are able to produce novel estimates for the zygoty, chorionicity, and amnionicity type frequencies of twins, triplets and quadruplets, which are presented in Section 3.3.

3.1 Estimates of Division Probabilities

As mentioned previously, no direct estimates for the probability that a zygote undergoes an early, middle, or late division exist in the literature. We give novel estimates based upon the known proportions of identical twins who are dichorionic/diamniotic (di/di), monochorionic/diamniotic (mono/di), and monochorionic/monoamniotic (mono/mono), which correspond to an early, middle, and late division, respectively. In particular, approximately 25% of twins are di/di, 73% are mono/di, and 2% are mono/mono [5, 8].

Now to leading order, the ratio of di/di identical twins to all identical twins is

$$\frac{\delta_e p_i^2 \beta_1^2}{\delta_e p_i^2 \beta_1^2 + \delta_m p_i \beta_2^2 + \delta_l p_i \beta_2^2 p_a^2} = 0.25,$$

the ratio of mono/di identical twins to all identical twins is

$$\frac{\delta_m p_i \beta_2^2}{\delta_e p_i^2 \beta_1^2 + \delta_m p_i \beta_2^2 + \delta_l p_i \beta_2^2 p_a^2} = 0.73,$$

and the ratio of mono/mono identical twins to all identical twins is

$$\frac{\delta_l p_i \beta_2^2 p_a^2}{\delta_e p_i^2 \beta_1^2 + \delta_m p_i \beta_2^2 + \delta_l p_i \beta_2^2 p_a^2} = 0.02.$$

Here all factors of the form $(1 - \delta_e)$, $(1 - \delta_m)$, or $(1 - \delta_l)$ are approximated to one for feasibility. Only two of these three equations are linearly independent, so we need an additional linearly independent equation in order to solve for the unknowns of δ_e , δ_m , and δ_l .

The additional linearly independent equation is produced by the ratio of identical twin sets to singletons, which can be derived from the Weinberg Differential Rule. As described in the Introduction, there are an estimated 39,624 sets of identical twins in the U.S. 1952–1954 birth data set, as well as 10,283,346 singletons, resulting in a ratio of 0.00385322. Now to leading order, the ratio of identical twins sets to singletons is

$$\frac{\delta_e p_i^2 \beta_1^2 + \delta_m p_i \beta_2^2 + \delta_l p_i \beta_2^2 p_a^2}{p_i \beta_1} = 0.00385322.$$

Solving the system of three linearly independent equations using basic linear algebra techniques results in $\delta_e \approx 0.001606$, $\delta_m \approx 0.003922$, and $\delta_l \approx 0.000430$. We observe that a middle division is the most likely, occurring about 0.4% of the time and is about twice as likely as an early division, which in turn is about four times as likely as a late division. These division probabilities, along with the other key parameter estimates, are summarized in Table 2.

Parameter	Notation	Estimate	Parameter	Notation	Estimate
anovulation	$P(O_0)$	0.007	survival with placenta shared by 1 embryo	β_1	0.8571
fertilization	p_f	0.3	survival with placenta shared by 2 embryos	β_2	0.7840
implantation	p_i	0.7	survival with placenta shared by 3 embryos	β_3	0.7696
early division	δ_e	0.001606	survival with placenta shared by 4 embryos	β_4	0.7544
middle division	δ_m	0.003922	relative survival for shared amniotic sac	p_a	0.5
late division	δ_l	0.000430			

Table 2: Summary of Key Parameter Estimates

3.2 Estimates of Polyovulation Probabilities

Using the birth type relative frequencies, combined with our fertilization, division and implantation, and miscarriage models, as well as the parameter values given in Table 2, we obtain the following estimates for polyovulation probabilities:

$$\begin{aligned} P(O_0) &\approx 0.007 \\ P(O_1) &\approx 0.957643 \\ P(O_2) &\approx 0.034784 \\ P(O_3) &\approx 0.000533 \\ P(O_4) &\approx 0.000040; \end{aligned}$$

hence, we estimate that more than one egg is released approximately 3.5% of the time. Note that these are estimates of population frequencies across fertile women in the U.S. An individual woman's polyovulation rates may differ from these population-level frequencies, and there is evidence that women who exhibit polyovulation in one cycle are more likely to exhibit polyovulation in future cycles as well. In addition, the polyovulation estimates were based upon Caucasian birth data, but there is evidence that polyovulation rates may vary across racial and ethnic lines [5].

3.3 Estimates of Zygosity, Chorionicity, and Amnionicity Probabilities

In Section 2.4, we defined $B_{k,z,c,a}$ to be the event that a pregnancy results in k live births with zygosity z , chorionicity c , and amnionicity a . By looking at the ratio of $B_{k,z,c,a}$ to B_k , the overall event that a pregnancy results in k live births, we obtain estimates of the zygosity, chorionicity, and amnionicity type relative frequencies for twins, triplets, and quadruplets. In particular, we estimate that for twins, 62.5% are dizygotic (i.e. fraternal)

and 37.5% are monozygotic (i.e. identical), which almost exactly matches the values from the Weinberg Differential Rule described in the Introduction. Of the monozygotic twins, we estimate that 25% are dichorionic/diamniotic, 73% are monochorionic/diamniotic, and 2% are monochorionic/monoamniotic. Note that these relative frequencies of identical twins are consistent with values from the literature [5, 8], which are based upon direct observation of the number of placentas and amniotic sacs from twin births. This is not surprising, since our estimates for the division probabilities were based upon these literature values.

Less is known in the literature about the zygosity, chorionicity, and amnionicity type relative frequencies for triplets and quadruplets since these higher order multiple births are fairly rare. Our novel estimates are given in Table 3. Note that for each possible zygosity, the most common outcome is for the chorionicity to equal the zygosity and for the amnionicity to equal the number of live births. This is an artifact of a middle division, which produces a shared placenta but separate amniotic sacs, having a significantly larger probability than an early or late division.

Now the work of Allen [1] does give estimates of the zygosity type relative frequencies for triplets and quadruplets, but does not give the added information regarding chorionicity and amnionicity. Allen estimates that 22% of triplets are monozygotic, 52% are dizygotic, and 26% are trizygotic, while for quadruplets, 12% are monozygotic, 35% are dizygotic, 27% are trizygotic, and 26% are quadrazygotic. These estimates were based upon extensions of the Weinberg Differential Rule to match the sex distributions of triplets and quadruplets.

The work of Guilherme et al. [4] analyzed the DNA and placentas of 64 sets of spontaneously conceived triplets to determine the precise zygosity and chorionicity relative frequencies within the data set, and found that 28% were

Triplets		Zygoty			Total
Chorionicity	Amnionicity	1	2	3	
1	1	0.000257			0.177807
	2	0.004759			
	3	0.172791			
2	2	0.001568	0.010443		0.449758
	3	0.057143	0.380604		
3	3	0.019547	0.130189	0.222699	0.372435
Total		0.256065	0.521236	0.222699	1

Quadruplets		Zygoty				Total
Chorionicity	Amnionicity	1	2	3	4	
1	1	0.000009				0.111614
	2	0.000147				
	3	0.003561				
	4	0.107897				
2	2	0.000051	0.000338			0.293947
	3	0.001216	0.008124			
	4	0.036994	0.247224			
3	3	0.000406	0.002716	0.004547		0.286505
	4	0.014774	0.098738	0.165324		
4	4	0.004210	0.028133	0.056527	0.219064	0.307934
Total		0.169265	0.385273	0.226398	0.219064	1

Table 3: Zygoty, Chorionicity, and Amnionicity Type Relative Frequencies for Triplets and Quadruplets

monozygoty, 50% were dizygoty, and 22% were trizygoty. Of the dizygoty triplets, 75% were dichoriony and 25% were trichoriony. Of the monozygoty triplets, 56% were monochoriony, 39% were dichoriony, and 5% were trichoriony. All of these values are fairly consistent with our estimates given in Table 3, and our work concurs with the assessment of Guilherme et al. that Allen’s estimates tend to slightly underestimate the proportion of multiples that are monozygoty. Moreover, we estimate that 2% of dizygoty triplets are diamnioty, while 0.1% of monozygoty triplets are monoamnioty and 2.5% are diamnioty. These amnionicity estimates are pertinent because shared amnioty sacs carry the highest risk of health complications.

4 Conclusion

In summary, we have analyzed the biological phenomenon of polyovulation and produced novel estimates for the probability of the human ovarian system ovulating k eggs in a single cycle, for $k \geq 0$. In our pursuit of these estimates, we utilized U.S. birth data from the 1950s, before the introduction of artificial reproductive technologies, and modeled the various stages that eggs undergo in order to reach live births, including fertilization, possible divi-

sion, implantation, and potential miscarriage. As a consequence of our model, we produced novel estimates for the probability that a zygote undergoes an early, middle, or late division, where the timing is defined by the number of days after fertilization. Furthermore, our model enabled us to provide estimates of the zygoty, chorionicity, and amnionicity type relative frequencies of twins, triplets, and quadruplets. While our main goal was to estimate polyovulation probabilities, these zygoty, chorionicity, and amnionicity estimates provide a valuable contribution in their own right to the study of multiple births.

Future work could include sensitivity analysis on the various literature estimates for anovulation, fertilization, implantation, and miscarriage, which were summarized in Table 2. In addition, the various assumptions made in our fertilization, division and implantation, and birth models could be reconsidered and generalized in order to have the models represent the true underlying biological processes to the fullest extent possible. In particular, we would like to generalize the division model to allow probabilities of secondary and tertiary divisions to differ from the probability of a primary division. It is biologically plausible that a zygote that has divided once may be either more or less susceptible to subsequent divisions, so both cases could be explored. We could also explore different implantation probabilities for individual versus

shared placentas, and our work here motivates the need for further biological research into how the underlying implantation mechanism might differ for a shared placenta compared to an individual placenta. In addition, the most unrealistic of our assumptions was that each embryo is miscarried independently of all other embryos; future work could remove the independence assumption and impart greater dependence for embryos sharing an amniotic sac or placenta.

Having estimated the probability of k eggs being ovulated in a single menstrual cycle, future work could seek to analyze the distribution of the number of eggs ovulated from the left versus right ovaries. One possibility is that the location of each egg released is independent of the locations of all other eggs released. This implies that the number of eggs released from the left ovary, given that there are a total of k eggs ovulated, follows a Binomial($k, \frac{1}{2}$) distribution, and likewise for the right ovary. A second possibility is that ipsilateral ovulation is favored where polyovulation tends to result from the overstimulation of a single ovary, and a third possibility is that contralateral ovulation is favored where the presence of the corpus luteum resulting from the ovulation of an egg from a particular ovary has an inhibitory effect on the release of additional eggs from the same ovary. All three of these cases are biologically reasonable [2], and there is very limited real data to test which of the three models is most likely. The work of Tong et al. [9] found that out of 27 women with spontaneous double ovulation, 7 had both eggs released from the left ovary, 14 had one egg released from each ovary, and 6 had both eggs released from the right ovary. This is consistent with a binomial distribution where the locations are independent, but there is need for more data and further analysis because it is possible that although double ovulation follows a binomial distribution, triple ovulation and quadruple ovulation may depart from a binomial model.

References

- [1] Allen, Gordon. "A differential method for estimation of type frequencies in triplets and quadruplets." *American journal of human genetics* 12.2 (1960): 210.
- [2] Baird, David T. "A model for follicular selection and ovulation: lessons from superovulation." *Journal of steroid biochemistry* 27.1 (1987): 15–23.
- [3] Blickstein, Isaac, and Louis G. Keith, eds. *Multiple pregnancy: epidemiology, gestation, and perinatal outcome*. CRC Press, 2005.
- [4] Guilherme, Romain, et al. "Zygosity and chorionicity in triplet pregnancies: new data." *Human Reproduction* 24.1 (2009): 100–105.
- [5] Hall, Judith G. "Twinning." *The Lancet* 362.9385 (2003): 735–743.
- [6] Malcolm, Christine E., and David C. Cumming. "Does anovulation exist in eumenorrheic women?." *Obstetrics & Gynecology* 102.2 (2003): 317–318.
- [7] Norwitz, Errol R., Danny J. Schust, and Susan J. Fisher. "Implantation and the survival of early pregnancy." *New England Journal of Medicine* 345.19 (2001): 1400–1408.
- [8] Shulman, Lee P., and John MG van Vugt, eds. *Prenatal medicine*. Taylor & Francis, 2006, p. 447.
- [9] Tong, Stephen, Simon Meagher, and Beverley Volenhoven. "Sonography: dizygotic twin survival in early pregnancy." *Nature* 416.6877 (2002): 142–142.
- [10] Weinberg, Wilhelm. "Beiträge zur physiologie und pathologie der mehrlingsgeburten beim menschen." *Archiv für die Gesamte Physiologie des Menschen und der Thiere* 88 (1901): 346–430.