A statistical evaluation of implantation after replacing one or more human embryos

D. E. Walters, R. G. Edwards* and M. L. Meistrich†

A.F.R.C. Statistics Group, Department of Applied Biology, University of Cambridge CB2 3DX, U.K., *Bourn Hall Clinic and Department of Physiology, University of Cambridge, Cambridge CB2 3EG, U.K. and †M. D. Anderson Medical Centre, Houston, Texas 77030, U.S.A.

Summary. All the available data relating to pregnancy and multiple implantations from Bourn Hall Clinic are used to assess the accuracy of two statistical models of implantation. The simple binomial model was found to be incapable of describing both the pregnancy rates and the incidence of multiple implantations, whereas a two-parameter model, incorporating both patient and embryo variability, provided a good fit to the data. The estimates obtained for the two parameters in the second model suggested that 36% of the embryos could implant, and that 43% of patients were capable of sustaining implantation. The Maximum Likelihood method of constructing 95% confidence limits on the estimates of the parameter values is demonstrated.

Introduction

In-vitro fertilization is being applied to the alleviation of human infertility with increasing success as various physiological and embryonic constraints are clarified. Implantation, however, remains a major obstacle. The chances of pregnancy improve progressively as 2 or more embryos are replaced, the increases conforming closely to the expected binomial probabilities on the assumption that each embryo has, independently, a constant probability of implantation (Edwards & Steptoe, 1983; Edwards, 1984). On some occasions when 3 embryos are replaced, implantation occurs in as many as 40% of all patients under 40 years of age. Many factors influence the chances of implantation, including the correct endocrine and uterine condition in the mothers, and the health and viability of the replaced embryos. There are obvious problems in studying human implantation, hence an attempt is made in this paper to quantify maternal and embryonic factors that influence implantation.

The number of multiple implantations when 2 or more embryos are replaced is often substantially greater than expected with the simple binomial model (Edwards, 1984). One possible explanation for this is that there is an interaction between the embryos in the uterus, the implantation of one embryo improving the probability of a second, and perhaps a third embryo, implanting. Although the simple binomial model may thus explain the incidence of pregnancies, it is inadequate to accommodate the frequency of multiple implantations. One reason could be that its single parameter fails to incorporate both embryo and patient variability. A two-parameter model may therefore be necessary to describe the entire process of implantation and multiple pregnancy.

In this paper all the available data relating to pregnancy and multiple implantations from Bourn Hall Clinic have been used to test both the simple binomial model and a more complex two-parameter model of human implantation after replacement of embryos into the uterus. The patients had received different hormonal stimulants with clomiphene and human menopausal gonadotrophin (HMG: Pergonal, Serono), and follicular maturation was induced by an endogenous LH surge or by an injection of 5000 i.u. hCG (Pregnyl, Organon). To illustrate the statistical methods, all patients have been combined into a single data set.
Data on implantation

Between 1 and 3 embryos were replaced into each patient, as described earlier (Edwards & Steptoe, 1983) with the incidence of pregnancy rising from approximately 15% to almost 40% with increasing number of embryos replaced. All short-lived biochemically suspected pregnancies were excluded from the data, only clinically-recognized pregnancies and abortions being included. Some pregnant patients were examined by ultrasound, and the number of implanted embryos was assessed from the ultrasound record if possible. Some patients were not seen during pregnancy since they returned to their own doctors soon after their embryos were replaced, and information on the number of children born was the only means of assessing the number of implanted embryos. Much data on abortions failed to provide the exact number of implanted fetuses and had to be excluded from detailed estimation of implantation.

The binomial model

If the probability of embryo implantation is \( p \) independently for each embryo, this parameter summarizing the entire variation in the system, the distributions of implantations when 1, 2 or 3 embryos are replaced are given by the binomial distribution, the appropriate expressions being displayed in the first column of Table 1. Clearly the probability of pregnancy is the complement of the probability of no implantations, and may be written quite simply as

\[
\text{Prob. (pregnancy)} = 1 - (1 - p)^k
\]

where \( k \) is the number of embryos replaced.

The accuracy of the binomial model may be tested by establishing whether the experimental data can be generated reasonably precisely by varying the single parameter \( p \).

<table>
<thead>
<tr>
<th>Number of embryos replaced</th>
<th>Number of embryos implanted</th>
<th>Single-parameter model ( p )</th>
<th>Two-parameter model ( w, p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>( (1 - p) = p_1 )</td>
<td>( 1 - wp = \theta_1 )</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>( p = p_2 )</td>
<td>( wp = \theta_2 )</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>( (1 - p)^2 = p_3 )</td>
<td>( w(1 - p)^2 + (1 - w) = \theta_3 )</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>( 2p(1 - p) = p_4 )</td>
<td>( 2wp(1 - p) = \theta_4 )</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>( p^2 = p_5 )</td>
<td>( wp^2 = \theta_5 )</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>( (1 - p)^3 = p_6 )</td>
<td>( w(1 - p)^3 + (1 - w) = \theta_6 )</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>( 3p(1 - p)^2 = p_7 )</td>
<td>( 3wp(1 - p)^2 = \theta_7 )</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>( 3p^2(1 - p) = p_8 )</td>
<td>( 3wp^2(1 - p) = \theta_8 )</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>( p^3 = p_9 )</td>
<td>( wp^3 = \theta_9 )</td>
</tr>
</tbody>
</table>

The parameters \( p_i \) and \( \theta_i \) have been defined merely to simplify the formulae in the text.

The two-parameter model

The previous model postulates that patient and embryo variability can be accommodated by a single parameter, representing the probability of embryo implantation. A more complex model may be proposed by suggesting that the probability of a patient being capable of implanting an embryo is \( w \), and that the probability of an embryo implanting is \( p \), independently for each embryo, with \( p \) also being independent of \( w \) (Speirs, Lopata, Gronow, Kellow & Johnstone, 1983). The
distribution of implantations when 1, 2 or 3 embryos are replaced is presented in the second column of Table 1. The probability of pregnancy when \( k \) embryos are replaced is given by

\[
\text{Prob. (pregnancy)} = 1 - [w(1 - p)^k + (1 - w)]
\]

This model was also fitted to the experimental data.

The statistical evaluation

The experimental data (two sets) which are used to test the models described above are displayed in Table 2. The upper part of the table shows the numbers of patients < 40 years of age who became pregnant when 1, 2 or 3 embryos were replaced. Only clinical pregnancies are included. The total available data on the number of implantations in patients with 2 or 3 replaced embryos are given in the lower part of the table, this information being unavailable for some of the patients, i.e. some who have aborted and others who are still pregnant or did not agree to follow-up studies. Each patient for whom implantation data were available is represented only once in Table 2.

Under the simple binomial model with single parameter \( p \), the joint probability (or Likelihood) of the observed configuration of pregnancies and multiple implantations may be derived from the formulae given in Table 1 and can be expressed quite simply.

\[
\text{Likelihood} = C(p_1)^{n_1} (p_2)^{n_2} (p_3)^{n_3} (1 - p_3)^{n_4} (p_6)^{n_5} (1 - p_6)^{n_6},
\]

\[
\times \left( \frac{p_4}{p_4 + p_5} \right)^{n_7} \left( \frac{p_5}{p_4 + p_5} \right)^{n_8}
\]

\[
\times \left( \frac{p_7}{1 - p_6} \right)^{n_9} \left( \frac{p_8}{1 - p_6} \right)^{n_{10}} \left( \frac{p_6}{1 - p_6} \right)^{n_{11}}
\]

where \( C \) is a function of the factorials of the observed frequencies, and the variables \( n_i \) are as defined in Table 2.

The value of \( p \) may be calculated, which maximizes the Likelihood of these data, and further, using asymptotic theory and the Likelihood Ratio Statistic, 95\% confidence limits may be estimated for \( p \). Also a test of hypothesis may be constructed on the adequacy of the single parameter to describe the configuration. The use of the Likelihood Function has been described in general terms by Cox & Hinkley (1974), and its application to binomial random variables by Cox (1970).

Adopting the 2-parameter model, and using the appropriate formulae given in Table 1, the joint probability of the configuration of pregnancies and multiple implantations is given by

\[
\text{Likelihood} = C(\theta_1)^{n_1} (\theta_2)^{n_2} (\theta_3)^{n_3} (1 - \theta_1)^{n_4} (\theta_6)^{n_5} (1 - \theta_6)^{n_6},
\]

\[
\times \left( \frac{\theta_4}{\theta_4 + \theta_5} \right)^{n_7} \left( \frac{\theta_5}{\theta_4 + \theta_5} \right)^{n_8}
\]

\[
\times \left( \frac{\theta_7}{1 - \theta_6} \right)^{n_9} \left( \frac{\theta_8}{1 - \theta_6} \right)^{n_{10}} \left( \frac{\theta_6}{1 - \theta_6} \right)^{n_{11}}
\]

\( C \) and \( n_i \) being defined as before.

Maximum Likelihood estimates may be derived for the pair of parameters and a 2-dimensional confidence region constructed, using methods similar to those applied for the single-parameter
Table 2. The observed frequencies of patients in the various categories, and the expected frequencies calculated from the models described in the text

<table>
<thead>
<tr>
<th>No. of embryos replaced</th>
<th>Outcome</th>
<th>Observed frequencies</th>
<th>Binomial model ((p = 0.148))</th>
<th>Binomial model ((p = 0.135))</th>
<th>Two-parameter model ((w = 0.43, p = 0.36))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pregnant</td>
<td>89 ((n_1))</td>
<td>86</td>
<td>79</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>Not pregnant</td>
<td>494 ((n_2))</td>
<td>497</td>
<td>504</td>
<td>493</td>
</tr>
<tr>
<td>2</td>
<td>Pregnant</td>
<td>121 ((n_3))</td>
<td>131</td>
<td>120</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>Not pregnant</td>
<td>357 ((n_4))</td>
<td>347</td>
<td>358</td>
<td>357</td>
</tr>
<tr>
<td>3</td>
<td>Pregnant</td>
<td>102 ((n_5))</td>
<td>120</td>
<td>111</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Not pregnant</td>
<td>213 ((n_6))</td>
<td>195</td>
<td>204</td>
<td>215</td>
</tr>
<tr>
<td>2</td>
<td>One implantation</td>
<td>60 ((n_7))</td>
<td>72</td>
<td>NF</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>Two implantations</td>
<td>18 ((n_8))</td>
<td>6</td>
<td>NF</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>One implantation</td>
<td>40 ((n_9))</td>
<td>52</td>
<td>NF</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Two implantations</td>
<td>16 ((n_{10}))</td>
<td>9</td>
<td>NF</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Three implantations</td>
<td>6 ((n_{11}))</td>
<td>1</td>
<td>NF</td>
<td>4</td>
</tr>
</tbody>
</table>

\(\chi^2 > 20\)
\(\chi^2 = 2.7\)
\(\chi^2 = 2.6\)
\(v = 5\)
\(v = 2\)
\(v = 4\)
\(P < 0.01\)
N.S.
N.S.

NF denotes frequencies not fitted by the model.
The \(\chi^2\) values were calculated from the Likelihood Ratio Statistic, and not from the expected frequencies listed above.
The symbols \((n_i)\) are used in the formulae quoted in the text.

model. A test of hypothesis of the adequacy of the model can be carried out, as for the previous model.

Results

Binomial model

The Maximum Likelihood estimate of the single parameter \(p\) was found to be 0.148, with estimated 95\% confidence limits of 0.134 and 0.163. The expected frequencies given in Table 2 demonstrate how this model was unable to explain satisfactorily all the features of the data. In particular, many more multiple implantations were obtained than would be predicted by this simple model. The Likelihood Ratio statistic which tests the adequacy of the model corresponded to a \(\chi^2\) value of more than 20, on 5 degrees of freedom, indicating a very poor fit.

This simple binomial model was also fitted to the data relating to pregnancies only, these results also being displayed in Table 2. The Maximum Likelihood estimate of \(p\) was now 0.135, with confidence limits of 0.122 and 0.150. The expected frequencies were generally quite close to those observed \((\chi^2 = 2.7)\), confirming that this model, although able to account for the incidence of pregnancies, was not able to accommodate the incidence of multiple implantations.

Two-parameter model

When the more complex, 2-parameter model was fitted to the data, the Maximum Likelihood estimates of the parameters were \(\hat{w} = 0.43\), and \(\hat{p} = 0.36\). The expected frequencies were now very close to those observed, demonstrating how this model is able to account for all the features of the
data (see Table 2). The construction of a confidence region requires a 2-dimensional diagram, which is shown in Text-fig. 1. The alignment of the confidence ellipse, along which run the lines of constant Likelihood, gives a good indication that the product of the 2 parameters \( w \times p \) is the important determinant in reproducing the data. Although the 2 parameters are necessary, much of the predictive power is contained in this product.

**Discussion**

The calculations described above demonstrate a method of assessing the parameters of embryo implantation in women. The data are obviously heterogeneous in origin, some having been obtained from scans in the first or second trimester, and the rest from birth records. Also, the data were derived from patients who had received different endocrinological treatments. The pooling of the data was carried out simply to demonstrate the fitting process, and to ensure a reasonably sized sample of patients with multiple implantations. Any insights into factors influencing human implantation would obviously be of considerable practical value in the clinical application of human in-vitro fertilization, especially the quantitative maternal and embryonic factors in implantation. A more detailed study of the effect of hormonal treatment will be published at a later date; suffice to state at this point that there was little evidence of systematic differences caused by the different forms of follicular stimulation, or the induction of follicular and oocyte maturation by an endogenous LH surge or the injection of hCG.

The values estimated for embryonic viability and for the ability of patients to implant an embryo will obviously be modified as data obtained from ultrasound scans accumulate because the incidence of multiple implantation is rather higher in early pregnancy (Edwards, 1984, 1985). Nevertheless, the observations show clearly that a simple binomial model, although being perfectly able to describe the pattern of pregnancies when more than one embryo is replaced, cannot accommodate the higher than expected number of multiple implantations. A subjective hypothesis to account for these discrepancies is that of ‘helping’, by which the probability of a second and perhaps of a third embryo implanting is considered to be greater than expected under the assumption of independence. In physiological terms, this hypothesis required that the implanting embryo is able to generate conditions conducive to further implantations. Unfortunately, the present data cannot establish whether this physiological situation actually occurs.

The more complex model has considerable appeal since its 2 parameters accommodate both patient and embryo variability. The model fits the data well, but only by setting the parameters at values which appear slightly unrealistic. The Maximum Likelihood solution estimates that only
43% of women can be brought to a condition conducive to implantation. It also estimates that as many as 36% of embryos are able to implant. The product \((w \times p)\) is 0.155, and is very close to the Maximum Likelihood estimate (0.148) of the single parameter in the binomial model.

However, this two-parameter model is a compromise between what is desirable and what is possible. The ideal model would contain a variable representing the maternal effect and another variable representing the embryo effect, with these variables following specified distributions in the population. It is very unlikely, for ethical as well as practical reasons, that data will become available to test such a model. In the present model it is assumed that a proportion of women \((w)\) in the population are capable of implantation at the appropriate time. The value of this parameter will obviously reflect a clinic's ability to prepare the patient so as to maximize her receptivity, as well as her innate ability to implant. The second parameter \(p\) defines the probability of embryo implantation as described above. These parameters are assumed to have fixed unknown values whose estimates are obtained from the data, and not, as seems more realistic, be distributed in some way in the population. Nevertheless despite this simplification, valuable information has been derived from a formulation of the model in these terms.

There are difficulties in the wholehearted acceptance of either of the two models to explain the observed incidence of implantation and multiple pregnancies. The simple model is unsatisfactory since it demands a second condition of embryo helping. On the other hand, the two-parameter model can fit the data only by setting the parameters at values which appear unlikely. There are no apparent endocrinological differences identifying the 43% of women capable of implantation from the remainder (Edwards et al., 1984), and most estimates of the probability of embryo implantation are far less than 0.36 (e.g. Speirs et al., 1983). Unfortunately, these authors also included the short-lived biochemical pregnancies in their calculations, which leads to difficulties in interpreting the conclusions.

A definitive choice between these 2 models, or indeed between other models as yet untried, could only be made after a carefully designed and conducted series of experiments. Since these experiments would almost certainly require the repeated use of a fertile and receptive patient, with different numbers of embryos being replaced on different occasions, it is most unlikely that such an experiment will be practical or ethical in clinics devoted to the cure of human infertility. An alternative approach might be to examine patients by ultrasound during implantation (Feichtinger & Kemeter, 1985), to establish whether embryos can implant successively, and so gain direct evidence of 'helping'.

In the assessment of various models it must be recognized that the extension of a model of a particular type with additional parameters will invariably improve the 'goodness of fit' to the data. A model's ability to reproduce the data cannot, in isolation, give it much credibility as a method of describing a biological process. Rather, the model should emerge initially from biological considerations, and also be capable of reproducing the numerical results. The main point to emerge from this investigation is that 2 parameters appear to be necessary to model the implantation process. The parameters defined above (in both models) are quite realistic, but of course are by no means the only ones which could be used. An alternative model based on the beta distribution (R. A. Elton, personal communication) provides an exceptional fit to these data and we hope to pursue the study of this and other models, using published data on implantation to assess their adaptability.

References


Received 5 October 1984