RESEARCH ARTICLE

Early embryo mortality in natural human reproduction: What the data say [version 1; peer review: 1 approved, 2 approved with reservations]

Gavin E. Jarvis

Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, CB2 3EG, UK

v1 **First published:** 25 Nov 2016, **5**:2765
https://doi.org/10.12688/f1000research.8937.1
Latest published: 07 Jun 2017, **5**:2765
https://doi.org/10.12688/f1000research.8937.2

Abstract

It is generally accepted that natural human embryo mortality during pregnancy is high – losses of 70% and higher from fertilisation to birth are frequently claimed. The first external sign of pregnancy occurs two weeks after fertilisation with a missed menstrual period. Establishing the fate of embryos before this is challenging, and hampered by a lack of data on the efficiency of fertilisation under natural conditions. Four distinct sources are cited to justify quantitative claims regarding embryo loss: (i) a hypothesis published by Roberts & Lowe in *The Lancet* is widely cited but has no practical quantitative value; (ii) life table analyses give consistent assessments of clinical pregnancy loss, but cannot illuminate losses at earlier stages of development; (iii) studies that measure human chorionic gonadotrophin (hCG) reveal losses in the second week of development and beyond, but not before; and (iv) the classic studies of Hertig and Rock offer the only direct insight into the fate of human embryos from fertilisation under natural conditions. Re-examination of Hertig's data demonstrates that his estimates for fertilisation rate and early embryo loss are highly imprecise and casts doubt on the validity of his numerical analysis. A recent re-analysis of hCG study data suggests that approximately 40-60% of embryos may be lost between fertilisation and birth, although this will vary substantially between individual women. In conclusion, it is clear that some published estimates of natural embryo mortality are exaggerated. Although available data do not provide a precise estimate, natural human embryo mortality is lower than is often claimed.

Keywords

early pregnancy loss , occult pregnancy , embryo mortality , human chorionic gonadotrophin , Hertig , pre-implantation embryo loss

Open Peer Review

Approval Status ✓ ✓ ?

	1	2	3
version 2 (revision) 07 Jun 2017	✓ view		
version 1 25 Nov 2016	? view	✓ view	? view

1. **Philippa Saunders**, University of Edinburgh, Edinburgh, UK
Douglas A Gibson, University of Edinburgh, Edinburgh, UK

2. **Jack Wilkinson**, University of Manchester, Manchester, UK

3. **Steven H. Orzack**, Fresh Pond Research Institute, Cambridge, USA
James E. Zuckerman, Harvard Medical School , Boston, USA

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Gavin E. Jarvis (gej1000@cam.ac.uk)

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

How to cite this article: Jarvis GE. **Early embryo mortality in natural human reproduction: What the data say [version 1; peer review: 1 approved, 2 approved with reservations]** F1000Research 2016, **5**:2765 https://doi.org/10.12688/f1000research.8937.1

First published: 25 Nov 2016, **5**:2765 https://doi.org/10.12688/f1000research.8937.1

Introduction

It is widely accepted that under natural circumstances, human embryo mortality is high, particularly immediately after fertilisation. Quantitative estimates of embryo loss are found in diverse media including television documentaries ("You made it through the first round" presented by Michael Mosley: video at http://www.bbc.co.uk/timelines/z84tsg8; transcript at http://a.files.bbci.co.uk/bam/live/content/z3b87hv/transcript: accessed on 22nd October, 2016), online educational videos ("Bill Nye: Can We Stop Telling Women What to Do With Their Bodies?" presented by Bill Nye, the Science Guy: video at https://www.youtube.com/watch?v=4IPrw0NYkMg: accessed on 22nd October, 2016), online museum exhibits ("Who Am I? What happens in week 1?" presented by The Science Museum; available at http://www.sciencemuseum.org.uk/WhoAmI/FindOutMore/Yourbody/Wheredidyoucomefrom/Howdoyougrowinthewomb/Whathappensinweek1: accessed on 22nd October, 2016), news reports ("Scientists get 'gene-editing' go-ahead" by James Gallagher: article at http://www.bbc.co.uk/news/health-35459054: accessed on 22nd October, 2016), as well as academic philosophical articles[1] and legal judgements[2]. Among reputable scientific publications, including medical and reproductive biology text books, scientific reviews and primary research articles, reported mortality estimates include: 30–70% before and during implantation[3]; >50%[4], 73%[5] and 80%[6] before the 6th week; 75% before the 8th week[7]; 70% in the first trimester[8]; 40–50% in the first 20 weeks[9]; and 49%[10], >50%[11,12], 53%[13], 54%[14], 60%[15], >60%[16], 63%[17,18], 70%[19–23], 50–75%[24], 76%[5,25], 78%[26], 80–85%[27], >85%[28], and 90%[29] total loss from fertilisation to term. The variance in these estimates is striking. 90% intrauterine mortality implies a maximal live birth fecundability of 10%, and only then if all other stages of the reproductive process are 100% efficient. Observed human fecundability is low compared to other animals[13], but at approximately 20–30%[4,30] it is still higher than implied by such a high embryo mortality rate.

Early human embryo mortality is of interest not only to reproductive biologists and fertility doctors, but also to ethicists[31], theologians[32] and lawyers[2]. Nevertheless, becoming pregnant and having children is of primary and personal importance to many women and their families. As with all biological processes, nothing works perfectly all the time[33], and failure to conceive and pregnancy loss are common problems. However, inconsistent estimates of early pregnancy loss are not reassuring, nor do they provide a sound basis for either a quantitative understanding of natural human reproductive biology or an unbiased appraisal of artificial reproductive technologies. The divergent and excessive values noted above therefore invite scrutiny of the evidence that supports them. In this article, I identify and re-evaluate published data that contribute to claims regarding natural human embryo mortality.

A quantitative framework for embryo mortality

A quantitative framework has been proposed to facilitate the calculation and comparison of embryo mortalities from fecundability and pregnancy loss data[34]. The model comprises conditional probabilities (π) of the following biological processes: (1) reproductive behaviours resulting in sperm-ovum-co-localisation per cycle = π_{SOC}; (2) successful fertilisation given sperm-ovum-co-localisation = π_{FERT}; (3) implantation of a fertilised ovum as indicated by increased levels of human chorionic gonadotrophin (hCG) = π_{HCG}; (4) progression of an implanted embryo to a clinically recognised pregnancy = π_{CLIN}; (5) survival of a clinical pregnancy to live birth = π_{LB}.

Fecundability is the probability of reproductive success per cycle, but may take different values depending on the definition of success. The following four fecundabilities broadly follow Leridon[30]:

1. Total (all fertilisations): $FEC_{TOT} = \pi_{SOC} \times \pi_{FERT}$
2. Detectable (implantation): $FEC_{HCG} = \pi_{SOC} \times \pi_{FERT} \times \pi_{HCG}$
3. Apparent (clinical): $FEC_{CLIN} = \pi_{SOC} \times \pi_{FERT} \times \pi_{HCG} \times \pi_{CLIN}$
4. Effective (live birth): $FEC_{LB} = \pi_{SOC} \times \pi_{FERT} \times \pi_{HCG} \times \pi_{CLIN} \times \pi_{LB}$

Hence, the probability that a fertilised egg will perish prior to implantation is $[1 - \pi_{HCG}]$, and prior to clinical recognition is $[1 - (\pi_{HCG} \times \pi_{CLIN})]$. In theory, embryonic mortality may be estimated at different stages; however, in practice, this depends on available data. Clinical and live birth fecundabilities are most easily quantified and most frequently reported. Total and detectable fecundabilities are less frequently reported, although of direct relevance.

What the data say

Publications containing data relevant to early human embryo mortality were identified primarily by tracing citations found in articles, reviews and textbooks. Systematic online searches did not capture all of these studies. Some are particularly old, many were not conducted to address the specific question, and others are in books or publications that are not adequately indexed. If not entirely complete, nevertheless the data presented form a substantial proportion of relevant, available scientific information on natural early human embryo mortality.

Studies that contribute analysis and data relevant to the quantification of natural human embryo mortality fall into the following four categories and will be considered in turn.

1. A speculative hypothesis published in *The Lancet*.
2. Life tables of intra-uterine mortality.
3. Studies of early pregnancy by biochemical detection of hCG.
4. Anatomical studies of Dr Arthur Hertig and Dr John Rock.

1. Where have all the conceptions gone?

In 1975, a short hypothesis published in *The Lancet* entitled "*Where Have All The Conceptions Gone?*" concluded that 78% of all conceptions were lost before birth[26]. It has been widely cited by both scientists[4,17,19,20,35] and non-scientists[36,37] alike. Conceptions among married women aged 20–29 in England and Wales in 1971 were estimated and compared to infants born in the same period. In this analysis (Table 1) there are reliable values, e.g., census data, and simple arithmetical calculations. However, speculative values

Table 1. Numerical estimates of conceptions and their loss in married women aged 20–29 in England and Wales in 1971. The table replicates the values and calculations of Roberts & Lowe[26] with more explanatory detail. In addition, it illustrates how introducing variance into speculative estimates influences the final calculated value of embryo loss. *Data type indicates whether the numerical value is reliable (e.g., derived from census data), the result of a simple arithmetical calculation, or speculative (shown in italics). §Values are the 2.5th and 97.5th percentile boundaries, assuming a normal distribution for the variables centred on Roberts & Lowe's values with a coefficient of variation of 20%. †Speculative values were adjusted either up or down by 25% compared to Roberts & Lowe's values. Values for 'Length of menstrual cycle' were adjusted by 10%. ‡The median values of the 2.5th and 97.5th percentile boundaries from 1,000 simulations, each containing 10,000 separate estimates for embryo loss. The derivation of these values is described in the text. Briefly, each separate estimate of embryo loss was calculated using variable speculative values that were obtained by random sampling from a normal distribution with a mean equal to the Roberts & Lowe value and a coefficient of variation of 20%. The median value of the mean percentage loss was 73.3% and of the median was 76.5%. ¥The most frequent duration of a menstrual cycle is 28 days but there is substantial variability and the mean length is generally 30–31 days[30].

Description of data	Data type*	Roberts & Lowe values	Low estimate values†	High estimate values†	95% data range (CV = 20%)§
Married women aged 20–29 in 1971	Reliable value	2,437,000	2,437,000	2,437,000	-
Frequency of coitus per married woman per week	*Speculative value*	*2*	*1.5*	*2.5*	*[1.2, 2.8]*
Weeks per year	Reliable value	52	52	52	-
Acts of coitus among married women per year	Calculation	253,448,000	190,086,000	316,810,000	-
Percentage of acts of coitus that are unprotected	*Speculative value*	*25%*	*19%*	*31%*	*[15%, 35%]*
Acts of unprotected coitus per year	Calculation	63,362,000	35,641,125	99,003,125	-
Length of menstrual cycle (days)	*Speculative value*	*28*	*31¥*	*25*	*[17, 39]*
Length of fertile period in each cycle (days)	*Speculative value*	*2*	*1.5*	*2.5*	*[1.2, 2.8]*
Acts of unprotected coitus during fertile period per year	Calculation	4,525,857	1,735,769	9,821,739	
Probability of fertilisation	*Speculative value*	*50%*	*38%*	*63%*	*[30%, 70%]*
Total fertilised ova per year	Calculation	2,262,929	650,913	6,138,587	-
Number of infants born (live and still) in 1971	Reliable value	505,000	505,000	505,000	-
Total number of lost embryos in 1971	Calculation	1,757,929	145,913	5,633,587	-
Percentage of embryos lost before live birth	**Calculation**	**78%**	**22%**	**92%**	**[37%, 90%]‡**

are necessary to perform the calculations. Three are biological: (1) fertilisation rate following unprotected coitus during the fertile period was estimated as 50% and supported by reference to Hertig[38] (although his estimate was 84%[33]); (2) the length of a menstrual cycle (28 days); and (3) the duration of the fertile period (2 days). These latter values are plausible, but also variable. No justification is provided for three behavioural variables: (1) coital frequency estimated at twice per week; (2) proportion of unprotected coital acts estimated at 25%; and (3) either a random or regular distribution of coital acts during menstrual cycles such that 1/14 of all coital acts fall within a fertile period.

The validity of Roberts & Lowe's conclusion depends largely on the accuracy and precision of these speculative values. The following two simple analyses illustrate the sensitivity of their conclusion on the speculative values.

1. When four of the speculative values are reduced by 25% (e.g., coital frequency reduced to 1.5/week) and cycle length increased by 10% (from 28 days to 31 days[30]), the estimate for embryo loss drops to 22%. The opposite operation (e.g., coital frequency increased to 2.5/week) results in an estimate of 92% (Table 1). Embryo loss of 22% is barely sufficient

to account for observed clinical losses, and 92% indicates a maximum FEC_{LB} of 8%. Neither scenario is biologically plausible.

2. A non-zero variance was applied to each speculative value reflecting their uncertain nature. Using the random number generator in Microsoft® Excel (Office 2010) simulated values were obtained by random sampling from normal distributions with means equal to Roberts & Lowe's speculative values with coefficients of variation equal to 20%. For simplicity, it was assumed that there was no covariance between the different speculative values. Table 1 shows the expected range within which 95% of these simulated values fall (e.g., coital frequency is 1.2–2.8/week). For each simulated record, a new estimate of embryo loss was calculated and from 10,000 of these, the mean, median and 2.5th and 97.5th percentiles of embryo loss were determined. This was repeated 1,000 times: the mean value of the simulated means was 73.3% and of the simulated medians was 76.5%. The mean values of the 2.5th and 97.5th percentile boundaries for embryo loss were 37% and 90% (Table 1). The same simulation was also performed using NONMEM 7.3.0® (Icon PLC, Dublin, Eire) and generated 100,000 data records. The outcome of this is shown in Figure 1. The code and simulated data values are in Dataset 1.

Dataset 1. Figure 1 data

http://dx.doi.org/10.5256/f1000research.8937.d140569

See README.docx for a description of the file.

The sole purpose of these simple sensitivity analyses is to illustrate that modest adjustments to Roberts & Lowe's original speculative values can result in any biologically plausible estimate for embryo loss. The output from the calculation is therefore substantially dependent on the subjectively selected input. Such an analysis has no practical quantitative value.

Other sources of bias in their model include the failure to account for intentionally terminated pregnancies and the reduced fecundability of already pregnant women and nursing mothers. Despite this, it was described as "*persuasive*"[39] and it has been claimed that "*it is still difficult to better the original calculations of Roberts and Lowe (1975)*"[19]. By contrast, others have noted that "*their calculations can be criticized*"[4] and are "*tenuous*"[40]. Considering its quantitative limitations, it has been cited surprisingly often[8,20,41].

2. Life tables of intrauterine mortality

Constructing a life table of intrauterine mortality is challenging since embryonic death may occur even before the presence of an embryo is recognised. Nevertheless, in 1977 Henri Leridon published a complete life table of intrauterine mortality[18]. Leridon highlighted the consequences of inappropriate analysis and the quantitative biases produced by alternative numerical methods. Overall, he discussed sixteen studies, and provided detailed commentary on six[42–47]. These data are summarised in Figure 2 and suggest that 12–24% embryos alive at 4 weeks' gestation (i.e., approx. 2 weeks' post-fertilisation) will perish before birth.

Dataset 2. Figure 2 data

http://dx.doi.org/10.5256/f1000research.8937.d140570

See README.docx for a description of the file.

Leridon described the Kauai Pregnancy Study[42] in particular detail. In this study, an attempt was made to identify every pregnancy on Kauai from 1953–56. Women were encouraged to enrol as soon as they missed a period. Early pregnancy loss may therefore have been overestimated, since not all amenorrhoea is caused by conception, although other studies that relied upon medically-identified pregnancies probably underestimated early pregnancy loss by not capturing all cases[48]. Whatever the truth, it is clear that, among the studies reviewed by Leridon, the Kauai Pregnancy Study revealed the highest levels of pregnancy loss (Figure 2).

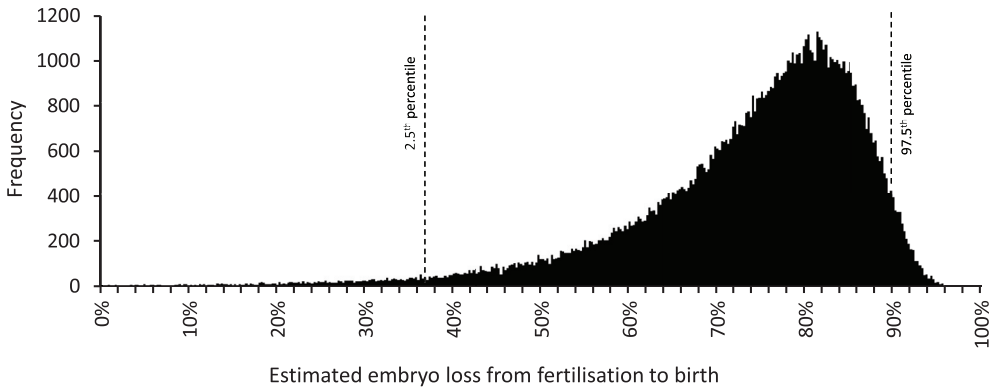

Figure 1. Distribution of embryo loss estimates from fertilisation to birth derived using a modified version of the model of Roberts & Lowe[26]. Embryo loss values were calculated using alternative speculative values (see text and Table 1) obtained by randomly sampling from normal distributions with mean values equal to the Roberts & Lowe's values with a coefficient of variation of 20%. 100,000 simulated embryo loss values were obtained. Frequencies within a bin size of 0.25% are shown. The 2.5th and 97.5th percentiles are indicated. The simulation was performed using NONMEM 7.3.0® (Icon PLC, Dublin, Eire). Simulated values are in Dataset 1.

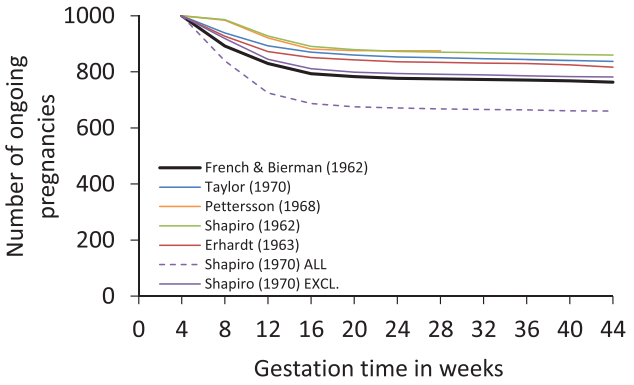

Figure 2. Graphical representation of the fate of 1,000 pregnancies in progress at 4 weeks' gestation (2 weeks' post-fertilisation). The figure is generated using values in Table 4.3 of Leridon[18] and are derived from six different studies (see text). The Kauai Pregnancy Study data[42] are shown in thick black. Data from Shapiro (1970)[46] were analysed either with all pregnancies included (ALL) or with those pregnancies excluded that aborted within one week of study entry (EXCL.). The greater loss observed with ALL may be due to a correlation between study entry and abortion risk. Based on these data, the risk of losing a pregnancy ongoing at 4 weeks' gestation ranges from 12.5% to 23.7% (excluding Shapiro (1970) ALL). Values are in Dataset 2.

All recorded pregnancies in the Kauai study were categorised by date of enrolment in four week intervals, beginning with 4–7 weeks' gestation. This time-staggered approach enabled risk of miscarriage to be associated with stage of gestation. However, despite considerable efforts, only 19% of the 3,197 recorded Kauai pregnancies were enrolled between 4–7 weeks' gestation, thereby reducing the precision of pregnancy loss estimates for this earliest of time intervals. Although pregnancies were grouped in four week periods, Leridon suggested that early mortality may change week by week, resulting in underestimation of pregnancy loss. He re-allocated the 592 study entries and 32 pregnancy losses for weeks 4–7 (Table 2) generating an overall probability of pregnancy loss during this period of 15.0%, higher than 10.8% originally reported[42]. Leridon's own description of this interpolation as "*risky*" can be illustrated by adjusting his re-allocation[18]. Transferring just two of the pregnancy losses out of or into the first week results in estimates of the 4–7 week pregnancy loss of 10.9% and 19.1% respectively (Table 2). The validity of adjusting Leridon's re-allocation may be questioned. However, pregnancy loss in week 4–5 of the Kauai Study would manifest as a menstrual period delayed by up to one week. This is far from being a robust pregnancy diagnosis and in different study[46], exclusion of pregnancy losses reported within one week of study entry resulted in substantially different loss probabilities (Figure 2) suggesting a confounding correlation between entry and loss[18]. Nevertheless, the re-allocation does reinforce a concern highlighted by Leridon, namely the uncertainty that affects the first probability. Clearly, these estimates of early loss should be treated with caution.

A more fundamental problem is that these data offer no insight into the fate of embryos prior to the earliest possible point of clinical pregnancy detection. Leridon completed his life table with values from Hertig's analysis[33]. He concluded that among 100 ova exposed to the risk of fertilisation, 16 are not fertilised, 15 die in week one (before implantation), and 27 die in week two (before the menstrual period). After two weeks his life table follows the Kauai probabilities closely ending with 31 live births. Leridon's table therefore indicates an embryo mortality of 50% (42/84) within the first two weeks after fertilisation and a total mortality of 63% (53/84) from fertilisation to birth.

Leridon's account of intrauterine mortality has been widely cited. However, its accuracy depends entirely on the quality and interpretation of the data from Hertig[33] and French & Bierman[42]. French & Bierman's approach probably resulted in an overestimate of total pregnancy loss and is certainly imprecise in its estimate of embryo loss in the four weeks following the first missed menstrual period. The reliability of Hertig's estimates of embryo loss in the two weeks following fertilisation is considered below.

Table 2. A speculative numerical re-allocation of entries and pregnancy losses during weeks 4–7 in the Kauai Pregnancy Study (KPS)[42]. Minor differences in the re-allocation of the earliest pregnancy losses have a substantial effect on the overall measure of pregnancy loss for that period. (Adapted from Table 4.2 in Leridon[18].)

Time period of gestation	New entries into study in each time period		Actual pregnancy losses in each time period		% pregnancy loss in each time period		Surviving pregnancies in each time period	
	Leridon's re-allocation	KPS	Leridon's re-allocation & [variants]	KPS	Leridon's re-allocation & [variants]	KPS	Leridon's re-allocation & [variants]	KPS
4–5	80	592	2 [0, 4]	32	5.0 [0.0, 10.0]	10.8	100 [100, 100]	100
5–6	120		6 [6, 6]		4.3 [4.3, 4.4]		95 [100, 90]	
6–7	180		10 [11, 9]		3.5 [3.9, 3.2]		91 [96, 86]	
7–8	212		14 [15, 13]		3.0 [3.2, 2.8]		88 [92, 83]	
4–8							85 [89, 81]	89.2
% loss							**15.0 [10.9, 19.1]**	**10.8**

3. Biochemical detection of pregnancy using hCG

Quantification of pregnancy loss requires pregnancy diagnosis. The earliest outward sign of pregnancy is a missed menstrual period, approximately 2 weeks after fertilisation, although amenorrhoea in women of reproductive age is not exclusively associated with fertilisation[49,50]. Several potentially diagnostic pregnancy-associated proteins have been identified[51] of which only one, Early Pregnancy Factor (EPF)[52], has been claimed to be produced by embryos within one day of fertilisation. However, there is doubt about the utility of EPF for diagnosing early pregnancy[53] and little has been published on it in the past five years.

Modern pregnancy tests detect human chorionic gonadotrophin (hCG), a highly glycosylated 37 kDa protein hormone produced by embryonic trophoblast cells[54]. Mid-cycle elevation of hCG is associated with embryo implantation[19,20,55]. Early assays for the detection of hCG were probably confounded by antibody cross-reactivity with luteinizing hormone[56] but modern tests are more specific and a positive result is a reliable indicator of early pregnancy. Highly sensitive assays have revealed low levels of hCG in non-pregnant women and healthy men[57]; hence, quantitative criteria are required to distinguish between non-pregnant women and those harbouring early embryos[55].

Figure 3 and Table 3 summarise findings from thirteen studies that used hCG to identify so-called early, occult or biochemical pregnancy loss, i.e., pregnancy loss between the initiation of implantation and clinical recognition[58–70]. Notwithstanding design and subject differences, estimates for clinical pregnancy loss, ranging from 8.3% - 21.2% (Figure 3), are similar to previous estimates (Figure 2). Estimates for early/occult loss ranged from 0% to 58.3% in studies[58–62] prior to Wilcox in 1988[63]. This high variance was probably due to reduced specificity and sensitivity of the hCG assays and sub-optimal study design[48,51,71–74]. Studies from 1988[63] onwards have produced more consistent data indicating early/occult loss of approximately 20% (Figure 3). In the three largest studies[63,66,70] pregnancies were clinically recognised only if they lasted ≥6 weeks after the onset of the last menstrual period[66,75]. Hence, early pregnancy losses in these studies included those lost up to approximately two weeks after a missed menstrual period: this may influence comparison of study results[34,73]. An overview of the thirteen studies suggests that overall pregnancy loss from first detection of hCG through to live birth is approximately one third (Table 3). This is consistent with another recent study which found that 98 out of 301 (32.6%) singleton pregnancies diagnosed by an early positive hCG test and followed-up to either birth or miscarriage were lost[76].

Dataset 3. Figure 3 data

http://dx.doi.org/10.5256/f1000research.8937.d140571

See README.docx for a description of the file.

The much cited Wilcox study[63] is the earliest of several large well-designed studies that made use of a specific and sensitive hCG assay and led to numerous further publications[75,77–83]. Two other studies (Zinaman[65] and Wang[66]) were similar in purpose, design and execution. These studies provide some of the best available data to calculate pregnancy loss between implantation and birth[34]. In each study, women intending to become pregnant and with no known fertility problems were recruited and hCG levels monitored

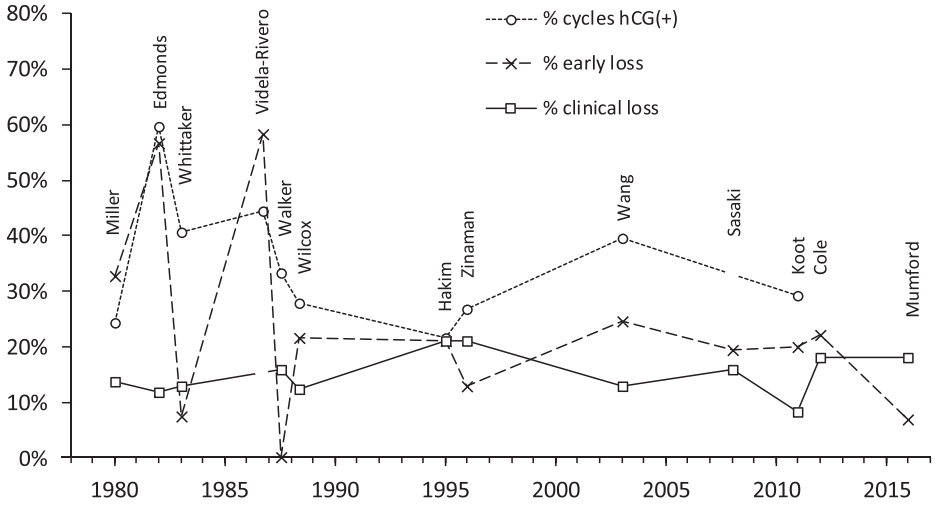

Figure 3. Summary of findings from thirteen studies that used hCG detection to diagnose early pregnancy. Data are arranged by publication date and the first author of the study is shown. Three datasets are shown: (i) the percentage of at risk reproductive cycles that were hCG positive; (ii) the percentage of hCG positive cycles that did not manifest as clinical pregnancies = early pregnancy loss; and (iii) the percentage of clinical pregnancies lost prior to 12 or 28 weeks or live birth (definitions vary between studies). A clinical pregnancy may be manifest by a missed period although criteria vary between studies. Videla-Rivero *et al.*[61], Sasaki *et al.*[67], Cole[69] and Mumford *et al.*[70] do not report sufficient data to calculate all three values. Values are in Dataset 3.

Table 3. Summary data from thirteen studies using hCG detection to diagnose pregnancy and identify early pregnancy loss. Raw FEC_{HCG} is the ratio of hCG pregnancies detected and the number of cycles monitored in each study. Where available, mean (SD) ages of the participating women are taken directly from the published study. In some cases mean and SD (indicated by *) or SD (indicated by †) were estimated based on published demographic characteristics. §These data relate to the whole study cohort (n=124) which included known sub-fertile women, and not just to the 74 apparently fertile women. ‡Mean value from Wilcox *et al.* (2001)[78]. ¶Some studies only provide data up to late pregnancy (e.g., up to 28 weeks) rather than to term. ND = no data. ¤Wilcox subsequently reported an additional hCG pregnancy which had not been detected and reported in the 1988 paper, making a total of 199 hCG pregnancies and 44 pre-clinical losses in the study group[75]. #Mumford reported data from aspirin- and placebo-treated subjects who had at least one prior miscarriage. Summary data from both treatment groups are included as there was no effect of aspirin[70].

First author	Year	Number of women	Age mean (SD) [range]	Number of cycles	hCG pregnancies detected	Raw FEC_{HCG}	Clinical pregnancies detected	% survival from hCG to clinical detection	% loss from hCG detection to live birth¶
Miller	1980	197	27 (4)*	623	152	24.4%	102	67.1%	42.4%
Edmonds	1982	82	27 (4)*	198	118	59.6%	51	43.2%	61.9%
Whittaker	1983	91	30 (3.7)†	226	92	40.7%	85	92.4%	19.6%
Videla-Rivero	1987	27	ND	27	12	44.4%	5	41.7%	ND
Walker	1988	38	27.4 [22–38]	75	25	33.3%	25	100%	16.0%
Wilcox	1988	221	30‡ (4)*	707	198¤	28.0%	155	78.3%	31.3%
Hakim	1995	74	31 (3)*§	305	66	21.6%	52	78.8%	37.9%
Zinaman	1996	200	30.6 (3.3)	432	116	26.9%	101	87.1%	31.3%
Wang	2003	518	24.9 (1.7)	1,561	618	39.6%	466	75.4%	35.7%
Sasaki	2008	110	[21–36]	ND	62	ND	50	80.6%	32.3%
Koot	2011	46	28.7 (3.3)	103	30	29.1%	24	80.0%	26.7%
Cole	2012	168	28.8 (4.4)	ND	127	ND	99	78.0%	36.2%
Mumford	2016	1088#	28.7 (4.8)	ND	785	ND	730	93.0%	23.9%

cycle by cycle in daily urine samples until they became pregnant. Most women were followed through to late pregnancy or birth. Although these studies provide evidence regarding the outcome of both clinical and hCG pregnancies, determining the fate of embryos prior to implantation is more difficult. To relate the study results to pre-implantation embryo loss, it is necessary to determine fecundability. In each study FEC_{CLIN} declined in successive cycles as the proportion of sub-fertile women increased. Hence, reported FEC_{HCG} values of 30%[65] and 40%[66], and FEC_{CLIN} values of 25%[63] and 30%[66] are biased underestimates of the fecundability of normal fertile women. A recent re-analysis of these data provides statistical evidence for discrete fertile and sub-fertile sub-cohorts within the study populations[34]. The proportions of sub-fertile women (mean [95% CI]) were estimated as 28.1% [20.6, 36.9] (Wilcox); 22.8% [12.9, 37.2] (Zinaman); and 6.0% [2.8, 12.3] (Wang). For normally fertile women, FEC_{HCG} was, respectively: 43.2% [35.6, 51.1]; 38.1% [32.7, 43.7]; and 46.2% [42.8, 49.6]. FEC_{CLIN} was: 33.9% [29.4, 38.6]; 33.3% [27.6, 39.6]; and 34.9% [33.0, 36.8]. There was no apparent difference in π_{CLIN} between fertile and sub-fertile sub-cohorts, which was estimated as: 78.3% [69.2, 85.3]; 87.5% [76.0, 93.9]; and 75.4% [71.5, 79.0][34].

Why do a proportion of menstrual cycles in women attempting to conceive fail to show any increase in hCG? Since $FEC_{HCG} = \pi_{SOC} \times \pi_{FERT} \times \pi_{HCG}$, there can be various causes for this failure including mistimed coitus, anovulation, failure of fertilisation or pre-implantation embryo death. Although FEC_{HCG} puts limits on the extent of pre-implantation embryo loss, uncertainty in the estimates of π_{SOC},π_{FERT} and π_{HCG} translates into uncertainty in estimates of pre-implantation embryo mortality. In the Wang study, for normally fertile women, FEC_{HCG} = 46.2%; hence, the absolute maximum value for pre-implantation embryo loss must be 53.8%, although only if $\pi_{SOC} = \pi_{FERT} = 1$, conditions both extreme and unlikely[34]. Studies of the relationship between coital frequency and conception indicate that fecundability is greater with daily compared to alternate day intercourse[34,84,85]. Hence, when coital frequency is less than once per day a proportion of reproductive failure will be due to mistimed coitus, i.e., $\pi_{SOC} < 1$. In the Wilcox study, coitus occurred on only 40% of the six pre-ovulatory days[34,79], and in the Zinaman study participants were advised that alternate day intercourse was optimal[65]. Based on the difference in fecundability between daily and alternate day intercourse as modelled by Schwartz[85], a value of $\pi_{SOC} = 0.80$ was used to calculate pre-implantation embryo mortality[34]. However, this is a speculative estimate, and in reality the value may be higher, or lower.

A further critical missing piece of the equation is knowledge of the efficiencies of fertilisation and implantation under normal, natural, propitious circumstances. Assuming that either of these

processes may be up to 90% efficient, and based on data from the three hCG studies[63,65,66], a plausible range for pre-implantation embryo loss in normally fertile women is 10–40% and for loss from fertilisation to birth, 40–60%[34]. Even with these wide ranges of mathematically possible outcomes, it is clear that estimates for total embryonic loss of 90%[29], 85%[28], 83%[31], 80–85%[6,27], 78%[26], 76%[5,25] and 70%[19–23] are excessive.

A previous review concluded that "*at least 73% of natural single conceptions have no real chance of surviving 6 weeks of gestation*"[5,86]. Live birth fecundability was estimated as "*not over 15%*", substantially lower than Leridon's 31%. Despite this discrepancy, Boklage's conclusions were derived from a review of data including several hCG studies[55,58–61,63] and Leridon's analysis[18]. He derived a model describing the survival probability of human embryos comprising the sum of two exponential functions:

$$P_t(\text{pregnancy survival}) = \mathbf{0.73}e^{-0.155t} + 0.27e^{-0.00042t}$$

...in which t is the time in days post-fertilization. This is the source of the 73% in the conclusion.

There are, however, serious problems with this analysis. Firstly, data presented as embryo survival probabilities at different times post-fertilization[55,58,59,61,63] are fecundabilities, i.e., successes per cycle, not per fertilised embryo. Secondly, for reasons that are unclear, data from Whittaker[60] and Leridon[18] were excluded from the modelling analysis and the data from an earlier Wilcox report[55] were included twice since this preliminary data had been incorporated into the later report[63]. Thirdly, the modelled data were normalised to a survival probability of 0.287 at 21 days post-fertilization. This value was derived from data published by Barrett & Marshall on the relationship between coital frequency and conception[84]. Barrett & Marshall had concluded that coitus during a single day alone, 2 days before ovulation resulted in a conception probability of 0.30. Boklage's value of 0.287 is his calculated equivalent. However, conception in this study was "*identified by the absence of menstruation, after ovulation*"[84]. Hence, 0.30 (and similarly, 0.287) is a clinical fecundability and not a measure of embryo survival. Furthermore, 0.30 is a non-maximal fecundability, since it was an estimate based on coitus on a single day (2 days before ovulation) within the cycle. Barrett & Marshall clearly report that as coital frequency increased so did the fecundability, up to a maximum of 0.68 associated with daily coitus[84].

Boklage's analysis can only make biological sense if it is assumed that every cycle in the Barrett & Marshall study resulted in fertilisation. Under these circumstances, failure to detect conception in 71.3% (1 – 0.287) of cycles would be due entirely to embryo mortality. However, this is highly implausible and explicitly contradicted by the higher estimate of fecundability reported[84]. Boklage's implicit assumption also contradicts his further conclusion that "*only 60–70% of all oocytes are successfully fertilized given optimum timing of natural insemination*"[5]. The vertical normalisation of the hCG study data to a value of 0.287 at 21 days is the principal determinant of the parameters that define the two exponential model. Any change in this value would commensurately alter the balance between the two implied sub-populations of embryos. Since it is evident that the value of 0.287 is neither an embryo survival rate nor even a maximal fecundability, it follows that quantitative conclusions from this analysis in relation to the survival of naturally conceived human embryos are of doubtful validity.

However, Boklage is right about two things: firstly, the difficulty of calculating pre-clinical losses, because "*In the place of the necessary numbers for the first few weeks of pregnancy we find editorially acceptable estimates which, while perhaps not far wrong, are difficult to defend with any precision*", and secondly, that the source of some of the only directly relevant data (even though he excluded it from his modelling analysis), namely, "*Hertig's sample is, and will probably remain, unique*".

4. The anatomical studies of Dr Arthur Hertig

At the start of the 1930s, no-one had ever seen a newly fertilised human embryo. It was barely 60 years since Oscar Hertwig had first observed fertilisation in sea urchins[87], and just 40 years before the birth of the first test tube baby[88,89]. In Boston, Dr Arthur Hertig and Dr John Rock's search to find early human embryos generated an irreplaceable collection which has left an indelible mark on our understanding of human embryology.

Hertig and Rock recruited 210 married women of proven fertility who presented for gynaecological surgery[38]. (In most of their publications, the number is given as 210[33,90,91] although 211 subjects are mentioned elsewhere[38].) Of these, 107 were considered optimal for finding an embryo because they apparently: (i) demonstrated ovulation; (ii) had at least one recorded coital date within 24 hours before or after the estimated time of ovulation; (iii) lacked pathologic conditions that would interfere with conception. Hertig examined the excised uteri and fallopian tubes, and over fifteen years found 34 human embryos aged up to 17 days[33,38,90–97]. Of these, 24 were normal and 10 abnormal[33,90]. (There is some confusion over this: in three publications[38,91,97], 21 embryos are described as normal and 13 as abnormal. It appears that the three alternatively described embryos (C-8299; C-8000; C-8290) were originally defined as abnormal based on their position or depth of implantation[38].) Table 4 provides information about the 34 embryos found in these 107 women. Although the study was primarily intended to find and describe early human embryos, Hertig subsequently used the data to derive estimates of reproductive efficiency including early embryo wastage[33,90].

Hertig's analysis[33,90] relies heavily on the 15 normal and 6 abnormal implanted embryos found in 36 women from cycle day 25 onwards. He assumed the 6 abnormal embryos would perish around the time of the first period concluding that fertility (% pregnant) at this stage = 42% (15/36). Of the 8 pre-implantation embryos identified (7 in the uterus and 1 in the fallopian tubes), 4 were abnormal. Hertig assumed the 4 normal embryos would implant successfully but that some of the abnormal ones would not, such that the proportion of normal embryos would increase from 50% (4/8) before implantation to 71% (15/21) after implantation as observed. Hence, among the 36 post-cycle day 25 cases, in addition to the 15 normal embryos, there must

Table 4. Summary of the characteristics of Hertig's 34 embryos (values are taken from Figure 4 in Hertig *et al.* (1959)). The embryos were collected from 107 out of 210 women. *In Hertig's figure, day 28 of the ovulatory cycle is identified with day 1 of the next cycle and is the day of the presumed missed period in cases where pregnancy had commenced. The 36 cases that provide the evidential foundation for his numerical analysis are shown in bold.

Day of cycle	Biological description/ stage	Approx. age of embryos (days)	Number of cases	Embryos found	Normal embryos	Abnormal embryos	Detection rate (%)
14	Ovulation ± fertilisation	0	0	0	0	0	
16–17	Embryo suspended in fallopian tube	2–3	9	1	1	0	11.1%
18–19	Embryo suspended in uterus	4–5	15	7	3	4	46.7%
20–24	Implantation	6–10	47	5	5	0	10.6%
25–3	**First missed period on day 28/1***	**11–16**	**36**	**21**	**15**	**6**	**58.3%**
Total			107	34	24	10	31.8%

have been 15 abnormal pre-implantation embryos of which 60% (9/15) failed to implant and were not observed, and 40% (6/15) did implant and were observed, although these 6 would have perished shortly afterwards. This left 6/36 eggs that must have been unfertilised. The ratio of 'unfertilised' : 'fertilised abnormal' : 'fertilised normal' was therefore 6:15:15, matching the 16% infertility (no fertilisation), 42% sterility (post-fertilisation death) and 42% fertility (reproductive success) reported in Figure 9 of Hertig's article, "*The Overall Problem in Man*"[33]. This is the source of Hertig's 84% fertilisation rate and 50% embryo loss before and during implantation, and is reproduced in Leridon's life table[18] as 84/100 eggs surviving at time zero (ovulation and fertilisation) and 42 surviving to 2 weeks (time of first missed period).

Hertig provides almost the entire body of evidence used to quantify natural human embryo loss in the first week post-fertilisation. Most claims regarding early human embryo mortality find their source here. Before considering how reliable the figures are, it is worth repeating Hertig's own caveat, namely, the lack of data on the efficiency of natural fertilisation[33]. All estimates of embryo mortality from fertilisation onwards are subject to commensurate inaccuracy in the absence of reliable fertilisation probabilities (i.e., π_{FERT}), which are "*surprisingly difficult to estimate*"[13].

There are several problems with Hertig's analysis. As noted by others, the observations are cross-sectional, but the inferences are longitudinal[48]. Hertig detected 21 embryos from 36 cases (58.3%) from cycle day 25 onwards. If this detection rate were representative, then on average, prior to day 25, the detection rate should either be the same or higher; however, they are all lower, and substantially so (Table 4). Hertig suggested that this was due to the technical difficulty of finding newly fertilised embryos. However, the detection rate for cycle days 18–19 was good (46.7%) and embryos one or two days younger would not have been much smaller, at which stage the detection rate was poor (11.1%). An alternative explanation for this discrepancy might simply be random variation. Furthermore, from cycle day 25 onwards, embryos would probably have produced hCG and therefore FEC_{HCG} would have been at least 58%. This is approximately double the equivalent values observed in more recent and robust hCG studies (Table 3) further suggesting that this subset of the data is not representative.

Despite having proven fertility, these women presented with gynaecological problems, suggesting sub-optimal reproductive function. Furthermore, Hertig's reproductively 'optimal' coital pattern does not include 2 days pre-ovulation and does include one day post-ovulation, conditions which are known not to maximise fertilisation[34,79,84,85,98]. Hence, detection rates before cycle day 25 may be more representative than those after. Given the numerical discrepancies, they cannot both be.

Hertig does not provide error estimates with his conclusions. In order to estimate the precision of his derived proportions, a bootstrap analysis was performed as follows: Hertig's 107 optimal cases were categorised according to stage of cycle (Category 1 = cycle days 16–19 (n=24); Category 2 = cycle days 20–24 (n=47); Category 3 = cycle days ≥25 (n=36)), and presence and type of embryos (Category 0 = no embryo (n=73); Category 1 = normal embryo (n=24); Category 3 = abnormal embryo (n=10)). Five hundred pseudo-datasets each containing 107 cases were generated using a balanced random re-sampling method using Microsoft Excel®. The original and pseudo datasets are in Dataset 4.

Dataset 4. Pseudo-datasets of Hertig's study, obtained via a bootstrap procedure

http://dx.doi.org/10.5256/f1000research.8937.d140572

See README.docx for a description of the files.

Hertig's numerical calculations, as detailed above, were repeated for each pseudo-dataset thereby generating 500 estimates for each parameter, from which median values and [95% CIs] were derived: fertility = 42% [26%, 59%]; sterility = 42% [5%, 182%]; infertility = 16% [-127%, 61%]; pre-implantation embryo survival probability = 69% [27%, 128%]; post-implantation to week two survival probability = 71% [50%, 91%]; detection rate for cycle day 25 onwards = 58% [41%, 74%]. Median values matched estimates calculated from the original dataset. Bootstrap 95% CIs for the day 25 detection rate (58%) matched those calculated using the "exact" method of Clopper & Pearson[99], [41%, 74%], which are a little wider than those calculated using the "more exact" method of Agresti & Coull[100], [42%, 73%]. (These analyses was performed using an online GraphPad® calculator accessed on 21st October 2016: http://www.graphpad.com/quickcalcs/ConfInterval1.cfm.) The congruence between these confidence intervals and the point estimates provides some reassurance that that the bootstrap procedure worked effectively. Estimates of parameters other than the day 25 detection rate (58%) are derived from more complex proportional relationships, and are therefore less precise. Table 5 reproduces a life table in the style of Leridon[18] and includes probabilities for each reproductive step with confidence intervals. These intervals (and some noted above) are impossibly wide highlighting further problems with Hertig's analysis.

Hertig's analysis omits 47 cases from cycle days 20–24, comprising 44% of his data. It is clear why he cannot use it, since all five embryos were normal and, given his mathematical and biological assumptions, five normal implanting embryos could not become 29% (6/21) abnormal post-implantation. Furthermore, the data that define the 50% proportion of abnormal pre-implantation embryos (i.e., 4/8) are so few that any numerical variation will make a substantial difference to derived proportions. If he had observed 3/8 abnormal embryos, his estimate of pre-implantation loss would have been 13% rather than 30%: for 5/8 it would have been 48%, with a fertilisation rate of 111%, which is clearly impossible. It seems therefore, that Hertig designed his analysis based on a post-hoc examination and selective use of the data. His own caveat about the lack of relevant and necessary data should be taken at least as seriously as his conclusions.

Hertig and Rock's contribution to human embryology is undeniable. However, their quantitative conclusions regarding early embryo mortality have a low precision that undermines their biological credibility or utility. Such estimates cannot be regarded as a reliable foundation upon which to evaluate and understand natural human reproduction.

Discussion

Answering the question "*How many fertilised human embryos die before or during implantation under natural conditions?*" is difficult. Relevant, credible data are in short supply. Among regularly cited publications, the *Lancet* hypothesis[26] is entirely speculative and in the view of the current author should cease to be used as an authoritative source. Clinical pregnancy studies are

Table 5. Life Table of egg survival and probabilities during the first two weeks of development derived solely from Hertig's data. The table is modelled on Leridon's life table[18] and includes his values for survivors and data from Hertig[33]. Probabilities are also shown for each stage of the early development process. Medians and 95% confidence intervals derived from a bootstrap analysis of Hertig's data indicate the precision in the estimates for fertilisation and embryo loss in the first two weeks. *Although Leridon's values are based on Hertig, they do not fully match. Leridon reports losses of 15 and 27 in the first and second weeks respectively. However, Hertig's 60% loss of abnormal pre-implantation embryos implies 25 (0.6 × 42) losses in the first week leaving 58, and 16 (58 × (6/21)) losses in the second week, leaving 42. ¥A value of π_{SOC} = 0.90 was used to avoid the calculation of probabilities greater than 1.

Week after Ovulation	Biological Description	Survivors (Leridon[18])	Survivors (Hertig[33])	Bootstrap Median [95% CIs]
	Number of Cycles	100	100	100 [100, 100]
0	Fertilised Eggs	84	83	84 [39, 227]
1	Implanted Embryos	69*	58	58 [41, 74]
2	Missed First Period	42	42	42 [26, 59]
Probabilities	**Biological Description**	**Probabilities**	**Probabilities**	**Bootstrap Median [95% CIs]**
$\pi_{SOC} \times \pi_{FERT}$	Fertilisation per cycle	0.84	0.83	0.84 [0.39, 2.27]
π_{FERT} (when π_{SOC} = 0.90¥)	Fertilisation per ideal insemination	0.93	0.93	0.93 [0.43, 2.52]
π_{HCG}	Fertilised egg implanting	0.82*	0.70	0.69 [0.27, 1.28]
π_{CLIN}	Implanted egg to clinical recognition	0.61*	0.71	0.71 [0.50, 0.91]
$\pi_{HCG} \times \pi_{CLIN}$	Fertilised egg to clinical recognition	0.50	0.50	0.50 [0.20, 0.88]

only useful for quantifying clinical pregnancy loss and contribute nothing to estimates of embryo mortality in the first two weeks' post-fertilisation. Even Hertig's unique dataset is inadequate to draw quantitative conclusions and oft-repeated values should be treated with scepticism. The hCG studies from 1988 onwards provide the best data for estimating embryo mortality although a lack of information on fertilisation rates[13,15,33,48,101] prevents satisfactory completion of the calculations. A recent re-analysis of these data has proposed plausible limits for reproductively normal women indicating that approximately 10–40% of embryos perish before implantation and 40–60% do so between fertilisation and birth[34]. However, these ranges are wide, particularly for pre-implantation mortality, reflecting the lack of appropriate data. Is there any possibility of narrowing down the numbers?

Two separate groups have previously collected embryos from women following carefully timed artificial insemination as part of fertility treatment. Insemination around the time of ovulation in women of proven fertility was followed 5 days later by uterine lavage to recover ova[102–105]. These data appear to hold promise for determining fertilisation efficiency and some authors have made quantitative inferences about embryo mortality from them[16,19,20]. However, such inferences are complicated by numerous confounding factors. For example, in one series[104], from 88 uterine lavages following artificial insemination by donor (AID), 4 unfertilised eggs, 6 fragmented eggs and 27 embryos from 2 cell to blastocyst stage were retrieved. In the 51 cycles in which no egg or embryo was retrieved, there was one retained pregnancy suggesting that the lavage and ova retrieval efficiency was reasonably high, albeit not perfect. These data therefore suggest that FEC_{TOT} was low (≈31/88 = 35%) although a proportion of fertilised eggs may have completely degenerated within the first 5 days. Assuming π_{SOC} was high (given the targeted insemination), this suggests that $\pi_{FERT} \approx 50\%$. In the context of the recent analysis[34], this implies that π_{HCG} is high and that levels of embryo mortality are therefore towards the lower end of the 10–40% and 40–60% ranges. However, the clinical pregnancy rate following transfer of the embryos was only 40%. This is equivalent to $\pi_{HCG} \times \pi_{CLIN}$. If $\pi_{CLIN} \approx 75\%$, as suggested by the hCG studies, this would mean that $\pi_{HCG} \approx 50\%$. This would imply that π_{FERT} is high, fertilised egg degeneration is high, occurs before day 5 and was therefore unobserved, and hence levels of embryo mortality tend towards the upper end of the 10–40% and 40–60% ranges.

It is possible that the lavage/transfer procedure reduced implantation and early developmental efficiency thereby reducing $\pi_{HCG} \times \pi_{CLIN}$. A comparison of AID pregnancy rates may provide some insight as suggested by the authors[104]. The clinical pregnancy rate in their pharmacologically unstimulated cohort was 12.5% (11/88) which is lower than an equivalent 18.9% observed for fresh semen AID[106], and also the live birth rate (which also incorporates clinical pregnancy losses) of 14.7% reported by the HFEA for AID in 2012 in unstimulated women aged 18–34[107]. These different success rates suggest that the lavage/transfer procedure did adversely affect implantation and early gestation with clear implications for quantitative extrapolation. Furthermore, the women who were embryo recipients were receiving fertility treatment and their overall fertility may have been lower than expected in a normal healthy cohort. In summary, it seems that there are too many unresolved variables in these data to narrow down estimates of fertilization (π_{FERT}) or implantation (π_{HCG}) rates.

With high fecundability, the range of possible embryo mortality rates falls. Red deer hinds have pregnancy rates of >85% following natural mating[108]: establishing numerical limits for embryo mortality under these efficient reproductive circumstances is more straightforward. By contrast, humans lack the instinct to mate predominantly during fertile periods thereby reducing observed reproductive efficiency substantially. In studies of early pregnancy loss, owing to sub-optimal coital frequency and cohorts including sub-fertile couples, natural fecundability was almost certainly not maximised[34]. Combining data on coital frequency and hCG elevation may help to address this. In a later analysis, applying the Schwartz model[85] to hCG data, Wilcox calculated a FEC_{HCG} value of 36% for high coital frequencies (>4 days with intercourse in 6 pre-ovulatory days)[79]. However, the model assumed that cycle viability was evenly distributed among couples, a condition which the authors recognised was not true and is contradicted by a subsequent analysis which suggests that approximately a quarter of the Wilcox cohort was sub-fertile[34]. If possible, focussing analytical attention on normally fertile women with the highest coital frequencies may help to further narrow the range of plausible embryo mortality.

In this review of natural early embryo mortality no use has been made of data from *in vitro* fertilisation (IVF) and associated laboratory studies. Sub-optimal conditions for embryo culture mean that it was[109,110] and probably still is[111] doubtful that reliable values can be extrapolated from laboratory *in vitro* to natural *in vivo* circumstances[20]. Importantly, the reproductive stages are also altered. In IVF, $\pi_{SOC} = 1$ and for transferred embryos $\pi_{FERT} = 1$. Furthermore, transferred embryos are selected based on quality criteria, however inexact those may be[111,112]. IVF program manipulations may reduce π_{HCG} compared to natural circumstances[3] and implantation failure remains a substantial issue for IVF[113,114]. Although for IVF cycles, the reported live birth rate per cycle has gone up (from 14% in 1991 to 25.4% in 2012[34]), comparison of IVF success rates and natural live birth fecundability values involves too many undefined variables to shed numerical light on early natural embryo development and mortality.

In vitro fertilisation *per se* may provide some insight into values of π_{FERT}, since $\pi_{SOC} = 1$, and successful fertilisation can be observed. In seven studies of natural cycle IVF, fertilisation was successful in 70.9% (443/625) of attempts[115–121]. If this represented natural, *in vivo* fertilisation, based on the recent analysis[34], it implies that $\pi_{HCG} \approx 0.75$, focusing estimates for pre-implantation embryo loss on 25%, and for total loss on 50%. However, high frequencies of chromosomal aberrations caused by the *in vitro* handling of human oocytes[122] can render any comparison of natural and assisted reproduction open to criticism[4].

In calculating summary values of embryo mortality, it is important to note that human fertility is as numerically heterogeneous as it could possibly be. Some couples are infertile and some are highly fertile. Excessive attention to averages and neglect of variances fosters a misleading appreciation of reality. The hCG studies clearly had both fertile and sub-fertile participants: use of overall values underestimated fecundability for the fertile majority[34]. Furthermore, apparently 'optimal' conditions for conception may not maximise human biological fecundability. Other biological factors also contribute to reproductive heterogeneity in humans; however, even after controlling for age-related decline, fecundability remains highly variable[107,123]. For intercourse occurring 2 days prior to ovulation, average fecundabilities resembled those previously published[124], but for couples at the 5th and 95th percentiles, fecundabilities were 5% and 83%. 83% fecundability implies a very low embryo mortality rate. In conclusion, apparent low fecundability in humans need not necessarily be caused by embryo mortality, but also defects of ovulation, mistimed coitus, or fertilisation failure[34]. Where fecundability is low, any or all of these factors may contribute.

Pregnancy loss and embryo mortality under natural conditions are real and substantial. However, estimates of 90%[29], 85%[28], 80%[6,27], 78%[26], 76%[5,25] and 70%[19–23] loss are excessive and not supported by available data. Estimates for clinical pregnancy loss are approximately 10–20%. For women of reproductive age, losses between implantation and clinical recognition are approximately 10–25%. Loss from implantation to birth is approximately one third[34,63,65,66].

Natural pre-implantation embryo loss remains quantitatively undefined. In the absence of knowledge of π_{SOC} and π_{FERT} it is almost impossible to estimate precisely. Hertig's estimate is 30%; however, mathematically and biologically implausible confidence intervals [-28%, 73%] betray the quantitative weaknesses in his data and analysis. The best available data are from studies monitoring daily hCG levels in women attempting to conceive[63,65,66]. Based on analyses of these data, in normal healthy women, 10–40% is a plausible range for pre-implantation embryo loss and overall pregnancy loss from fertilisation to birth is approximately 40–60%[34]. This latter range is similar to, although a little narrower than the 25–70% suggested by Professor Robert Edwards[125].

In the absence of suitable data to quantify pre-implantation loss, many published articles and reviews merely restate previously published values[6,20,21]. It has been suggested that "*for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias*"[126]. Widely held views on early embryo mortality may reflect an entrenched and biased view of the biology. For example, the Macklon "Black Box" review[20] has been cited over 200 times (Web of Knowledge citations on 10th October 2016) with many articles explicitly referencing its 30% survival/70% failure value[8,21,113,127–133]. Macklon's quantitative summary in his "Pregnancy Loss Iceberg" (30% implantation failure; 30% early pregnancy loss; 10% clinical miscarriage; 30% live births) is a direct, unedited reproduction of estimates published over 10 years previously[19]. 30% pre-implantation loss fairly represents Hertig's conclusions although, as has been shown, this estimate is highly imprecise. However, Macklon misrepresents the best data which he reviews[63,65]. Wilcox reports early pregnancy loss (i.e., [1 - π_{CLIN}]) of 21.7% whereas Macklon's iceberg implies that 43% (30/70) of implanting embryos fail before clinical recognition. The iceberg's clinical loss rate of 25% (10/40) is also higher than relevant data indicate (Figure 2 & Figure 3). Total loss of implanting (hCG+) embryos (i.e., [1 - (π_{CLIN} × π_{LB}]) is 57% (40/70) according to the iceberg. By contrast, Wilcox[63] and Zinaman[65], both included in Macklon's review, both report that only 31% of hCG positive pregnancies fail.

If Macklon's (and Chard's[19]) estimates are excessive as the data suggest, this casts doubt on claims[113,132] that the frequency of embryonic abnormalities observed *in vitro* is representative of the natural *in vivo* situation. In turn, this implies that many of the chromosomal abnormalities observed in *in vitro* human embryos are, to a greater extent than currently recognised[113], an artefact of the clinical and experimental context of assisted reproduction technologies.

In attempting to quantify pre-implantation embryo mortality it is easy to appreciate why "*a claim of 'no significant difference' might easily be sustained against any interpretation proffered*"[48], and why estimates are "*difficult to defend with any precision*"[5]. In conclusion, "*poor estimates of fertilization failure rate and the mortality at 2 weeks after fertilisation*"[15] drawn "*from unusual or biased samples*"[134] indicate that the "black box" of early pregnancy loss[20] is not as wide open as has been thought.

Data availability

F1000Research: Dataset 1. Figure 1 data, 10.5256/f1000research.8937.d140569[135]

F1000Research: Dataset 2. Figure 2 data, 10.5256/f1000research.8937.d140570[136]

F1000Research: Dataset 3. Figure 3 data, 10.5256/f1000research.8937.d140571[137]

F1000Research: Dataset 4. Pseudo-datasets of Hertig's study, obtained via a bootstrap procedure, 10.5256/f1000research.8937.d140572[138]

Competing interests

No competing interests were disclosed.

Grant information

The author(s) declared that no grants were involved in supporting this work.

Acknowledgements

Thanks are due to Professor David Paton, Dr Paul Schofield and Dr Amanda Sferruzzi-Perri for reviewing and providing helpful comments during the writing of this article.

References

1. Ord T: **The scourge: moral implications of natural embryo loss.** *Am J Bioeth.* 2008; **8**(7): 12–9.
PubMed Abstract | Publisher Full Text
2. **R (on the application of Smeaton) v Secretary of State for Health.** [2002] EWHC 610 (Admin), [2002] All ER (D) 115 (Apr), 2002.
Reference Source
3. Kennedy TG: **Physiology of implantation.** *In Vitro Fert Ass Rep.* 1997; 729–35.
Reference Source
4. Benagiano G, Farris M, Grudzinskas G: **Fate of fertilized human oocytes.** *Reprod Biomed Online.* 2010; **21**(6): 732–41.
PubMed Abstract | Publisher Full Text
5. Boklage CE: **Survival probability of human conceptions from fertilization to term.** *Int J Fertil.* 1990; **35**(2): 75, 79–80, 81–94.
PubMed Abstract
6. Vitzthum VJ, Spielvogel H, Thornburg J, *et al.*: **A prospective study of early pregnancy loss in humans.** *Fertil Steril.* 2006; **86**(2): 373–9.
PubMed Abstract | Publisher Full Text
7. Bainbridge DR: **Making Babies. A Visitor Within**. London: Phoenix; 2001; 101–62 at 59ff.
Reference Source
8. Ramos-Medina R, García-Segovia Á, León JA, *et al.*: **New decision-tree model for defining the risk of reproductive failure.** *Am J Reprod Immunol.* 2013; **70**(1): 59–68.
PubMed Abstract | Publisher Full Text
9. Norwitz ER, Schust DJ, Fisher SJ: **Implantation and the survival of early pregnancy.** *N Engl J Med.* 2001; **345**(19): 1400–8.
PubMed Abstract | Publisher Full Text
10. James WH: **The incidence of spontaneous abortion.** *Popul Stud (Camb).* 1970; **24**(2): 241–5.
PubMed Abstract | Publisher Full Text
11. Silver RM, Branch DW: **Sporadic and recurrent pregnancy loss.** In: Reece EA, Hobbins JC, editors. *Clinical Obstetrics: The Fetus and Mother.* 3rd ed: Blackwell Publishing; 2007; 143–60.
Publisher Full Text
12. Nishimura H: **Fate of human fertilized eggs during prenatal life: present status of knowledge.** *Okajimas Folia Anat Jpn.* 1970; **46**(6): 297–305.
PubMed Abstract | Publisher Full Text
13. Short RV: **When a conception fails to become a pregnancy.** *Ciba Found Symp.* 1978; (64): 377–94.
PubMed Abstract
14. Opitz JM: **The Farber lecture. Prenatal and perinatal death: the future of developmental pathology.** *Pediatr Pathol.* 1987; **7**(4): 363–94.
PubMed Abstract | Publisher Full Text
15. Biggers JD: **Risks of *In Vitro* Fertilization and Embryo Transfer in Humans.** In: Crosignani PG, Rubin BL, editors. *In Vitro Fertilization and Embryo Transfer.* London: Academic Press; 1983; 393–410.
Reference Source
16. Johnson MH: **Chapter 15: Fetal Challenges.** *Essential Reproduction.* 7th ed. Oxford: Wiley-Blackwell; 2013; 258–69.
Reference Source
17. Biggers JD: ***In vitro* fertilization and embryo transfer in human beings.** *N Engl J Med.* 1981; **304**(6): 336–42.
PubMed Abstract | Publisher Full Text
18. Leridon H: **Intrauterine Mortality.** *Human Fertility: The Basic Components.* Chicago: The University of Chicago Press; 1977; 48–81.
Reference Source
19. Chard T: **Frequency of implantation and early pregnancy loss in natural cycles.** *Baillieres Clin Obstet Gynaecol.* 1991; **5**(1): 179–89.
PubMed Abstract | Publisher Full Text
20. Macklon NS, Geraedts JP, Fauser BC: **Conception to ongoing pregnancy: the 'black box' of early pregnancy loss.** *Hum Reprod Update.* 2002; **8**(4): 333–43.
PubMed Abstract | Publisher Full Text
21. Ford HB, Schust DJ: **Recurrent pregnancy loss: etiology, diagnosis, and therapy.** *Rev Obstet Gynecol.* 2009; **2**(2): 76–83.
PubMed Abstract | Free Full Text
22. McCoy RC, Demko Z, Ryan A, *et al.*: **Common variants spanning *PLK4* are associated with mitotic-origin aneuploidy in human embryos.** *Science.* 2015; **348**(6231): 235–8.
PubMed Abstract | Publisher Full Text
23. Loke YW, King A: **Human Implantation: Cell Biology and Immunology.** Cambridge: Cambridge University Press; 1995.
Reference Source
24. American College of Obstetricians and Gynecologists: **Technical Bulletin No. 212: Early pregnancy loss.** *Int J Gynaecol Obstet.* 1995; **51**(3): 278–85.
PubMed Abstract | Publisher Full Text
25. Drife JO: **What proportion of pregnancies are spontaneously aborted?** *Brit Med J.* 1983; **286**(6361): 294.
26. Roberts CJ, Lowe CR: **Where have all the conceptions gone?** *Lancet.* 1975; **305**: 498–9.
Publisher Full Text
27. Johnson MH, Everitt BJ: **Chapter 15: Fertility.** *Essential Reproduction.* 5th ed. Oxford: Wiley-Blackwell; 2000; 251–74.
Reference Source
28. Braude PR, Johnson MH: **The Embryo in Contemporary Medical Science.** In: Dunstan GR, editor. *The Human Embryo: Aristotle and the Arabic and European Traditions.* Exeter: University of Exeter Press; 1990; 208–21.
Reference Source
29. Opitz JM: **Human Development - The Long and the Short of it.** In: Furton EJ, Mitchell LA, editors. *What is Man, O Lord? The Human Person in a Biotech Age; Eighteenth Workshop for Bishops.* Boston, MA: The National Catholic Bioethics Center; 2002; 131–53.
30. Leridon H: **Fecundability.** *Human Fertility: The Basic Components.* Chicago: The University of Chicago Press; 1977; 22–47.
Reference Source
31. Harris J: **Stem cells, sex, and procreation.** *Camb Q Healthc Ethics.* 2003; **12**(4): 353–71.
PubMed Abstract | Publisher Full Text
32. Rahner K: **Theological Investigations, Vol IX.** London: DLT; 1972.
33. Hertig AT: **The Overall Problem in Man**. In: Benirschke K, editor. *Comparative Aspects of Reproductive Failure.* An International Conference at Dartmouth Medical School. Berlin: Springer Verlag; 1967.
Publisher Full Text
34. Jarvis GE: **Estimating limits for natural human embryo mortality [version 1; referees: 2 approved].** *F1000Res.* 2016; **5**: 2083.
Publisher Full Text
35. Bulletti C, Flamigni C, Giacomucci E: **Reproductive failure due to spontaneous abortion and recurrent miscarriage.** *Hum Reprod Update.* 1996; **2**(2): 118–36.
PubMed Abstract | Publisher Full Text
36. Devolder K, Harris J: **The ambiguity of the embryo: Ethical inconsistency in the human embryonic stem cell debate.** *Metaphilosophy.* 2007; **38**(2–3): 153–69.
Publisher Full Text
37. Green RM: **The Human Embryo Research Debates: Bioethics in the Vortex of Controversy.** Oxford: Oxford University Press; 2001.
Reference Source
38. Hertig AT, Rock J, Adams EC: **A description of 34 human ova within the first 17 days of development.** *Am J Anat.* 1956; **98**(3): 435–93.
PubMed Abstract | Publisher Full Text
39. **Letter: Where have all the conceptions gone?** *Lancet.* 1975; **1**(7907): 636–7.
PubMed Abstract | Publisher Full Text
40. Cooke ID: **Failure of implantation and its relevance to subfertility.** *J Reprod Fertil Suppl.* 1988; **36**: 155–9.
PubMed Abstract
41. Catalano RA, Saxton KB, Bruckner TA, *et al.*: **Hormonal evidence supports the theory of selection *in utero*.** *Am J Hum Biol.* 2012; **24**(4): 526–32.
PubMed Abstract | Publisher Full Text | Free Full Text
42. French FE, Bierman JM: **Probabilities of fetal mortality.** *Public Health Rep.* 1962; **77**(10): 835–47.
PubMed Abstract | Free Full Text
43. Shapiro S, Jones EW, Densen PM: **A life table of pregnancy terminations and correlates of fetal loss.** *Milbank Mem Fund Q.* 1962; **40**(1): 7–45.
PubMed Abstract
44. Erhardt CL: **Pregnancy Losses in New York City, 1960.** *Am J Public Health Nations Health.* 1963; **53**(9): 1337–52.
PubMed Abstract | Publisher Full Text | Free Full Text
45. Pettersson F: **Epidemiology of Early Pregnancy Wastage.** Stockholm: Svenska Bokförlaget; 1968.
Reference Source
46. Shapiro S, Levine HS, Abramowicz M: **Factors associated with early and late fetal loss.** *Adv Planned Parenthood.* 1970; **6**: 45–63.
47. Taylor WF: **The Probability of Fetal Death**. In: Fraser FC, McCusick VA, editors. *Congenital Malformations.* Amsterdam: Excerpta Medica; 1970; 307–20.
48. Kline J, Stein Z, Susser M: **Conception and Reproductive Loss: Probabilities.** *Conception to Birth. Epidemiology of Prenatal Development.* New York: OUP; 1989; 43–68.
49. Master-Hunter T, Heiman DL: **Amenorrhea: evaluation and treatment.** *Am Fam Physician.* 2006; **73**(8): 1374–82.
PubMed Abstract
50. Committee on Practice Bulletins—Gynecology: **Practice bulletin no. 128: diagnosis of abnormal uterine bleeding in reproductive-aged women.** *Obstet Gynecol.* 2012; **120**(1): 197–206.
PubMed Abstract | Publisher Full Text
51. Grudzinskas JG, Nysenbaum AM: **Failure of human pregnancy after implantation.** *Ann N Y Acad Sci.* 1985; **442**: 38–44.
PubMed Abstract | Publisher Full Text
52. Morton H, Rolfe B, Clunie GJ: **An early pregnancy factor detected in human serum by the rosette inhibition test.** *Lancet.* 1977; **1**(8008): 394–7.
PubMed Abstract | Publisher Full Text
53. Chard T, Grudzinskas JG: **Early pregnancy factor.** *Biol Res Pregnancy Perinatol.* 1987; **8**(2 2D Half): 53–6.
PubMed Abstract

54. Cole LA: **hCG, the wonder of today's science.** *Reprod Biol Endocrinol.* 2012; **10**: 24.
PubMed Abstract | Publisher Full Text | Free Full Text

55. Wilcox AJ, Weinberg CR, Wehmann RE, *et al.*: **Measuring early pregnancy loss: laboratory and field methods.** *Fertil Steril.* 1985; **44**(3): 366–74.
PubMed Abstract

56. Regan L: **A prospective study of spontaneous abortion.** In: Beard RW, Sharp F, editors. *Early Pregnancy Loss: Mechanisms and Treatment.* Springer-Verlag; 1988; 23–37.
Publisher Full Text

57. Odell WD, Griffin J: **Pulsatile secretion of human chorionic gonadotropin in normal adults.** *N Engl J Med.* 1987; **317**(27): 1688–91.
PubMed Abstract | Publisher Full Text

58. Miller JF, Williamson E, Glue J, *et al.*: **Fetal loss after implantation. A prospective study.** *Lancet.* 1980; **2**(8194): 554–6.
PubMed Abstract | Publisher Full Text

59. Edmonds DK, Lindsay KS, Miller JF, *et al.*: **Early embryonic mortality in women.** *Fertil Steril.* 1982; **38**(4): 447–53.
PubMed Abstract | Publisher Full Text

60. Whittaker PG, Taylor A, Lind T: **Unsuspected pregnancy loss in healthy women.** *Lancet.* 1983; **1**(8334): 1126–7.
PubMed Abstract | Publisher Full Text

61. Videla-Rivero L, Etchepareborda JJ, Kesseru E: **Early chorionic activity in women bearing inert IUD, copper IUD and levonorgestrel-releasing IUD.** *Contraception.* 1987; **36**(2): 217–26.
PubMed Abstract | Publisher Full Text

62. Walker EM, Lewis M, Cooper W, *et al.*: **Occult biochemical pregnancy: fact or fiction?** *Br J Obstet Gynaecol.* 1988; **95**(7): 659–63.
PubMed Abstract | Publisher Full Text

63. Wilcox AJ, Weinberg CR, O'Connor JF, *et al.*: **Incidence of early loss of pregnancy.** *N Engl J Med.* 1988; **319**(4): 189–94.
PubMed Abstract | Publisher Full Text

64. Hakim RB, Gray RH, Zacur H: **Infertility and early pregnancy loss.** *Am J Obstet Gynecol.* 1995; **172**(5): 1510–7.
PubMed Abstract | Publisher Full Text

65. Zinaman MJ, Clegg ED, Brown CC, *et al.*: **Estimates of human fertility and pregnancy loss.** *Fertil Steril.* 1996; **65**(3): 503–9.
PubMed Abstract | Publisher Full Text

66. Wang X, Chen C, Wang L, *et al.*: **Conception, early pregnancy loss, and time to clinical pregnancy: a population-based prospective study.** *Fertil Steril.* 2003; **79**(3): 577–84.
PubMed Abstract | Publisher Full Text

67. Sasaki Y, Ladner DG, Cole LA: **Hyperglycosylated human chorionic gonadotropin and the source of pregnancy failures.** *Fertil Steril.* 2008; **89**(6): 1781–6.
PubMed Abstract | Publisher Full Text

68. Koot YE, Boomsma CM, Eijkemans MJ, *et al.*: **Recurrent pre-clinical pregnancy loss is unlikely to be a 'cause' of unexplained infertility.** *Hum Reprod.* 2011; **26**(10): 2636–41.
PubMed Abstract | Publisher Full Text

69. Cole LA: **Hyperglycosylated hCG and pregnancy failures.** *J Reprod Immunol.* 2012; **93**(2): 119–22.
PubMed Abstract | Publisher Full Text

70. Mumford SL, Silver RM, Sjaarda LA, *et al.*: **Expanded findings from a randomized controlled trial of preconception low-dose aspirin and pregnancy loss.** *Hum Reprod.* 2016; **31**(3): 657–65.
PubMed Abstract | Publisher Full Text

71. Wilcox AJ, Baird DD, Weinberg CR, *et al.*: **The use of biochemical assays in epidemiologic studies of reproduction.** *Environ Health Perspect.* 1987; **75**: 29–35.
PubMed Abstract | Publisher Full Text | Free Full Text

72. Brattebø G: **Occult biochemical pregnancy: fact or fiction?** *Br J Obstet Gynaecol.* 1989; **96**(2): 252–4.
PubMed Abstract | Publisher Full Text

73. Walker EM, Lewis M, Howie PW: **Authors' reply.** *Br J Obstet Gynaecol.* 1989; **96**(2): 253–4.
Publisher Full Text

74. Wilcox AJ, Weinberg CR, Baird DD: **Subclinical embryonic loss.** *Fertil Steril.* 1989; **51**(5): 907–8.
PubMed Abstract | Publisher Full Text

75. Wilcox AJ, Weinberg CR, Baird DD: **Risk factors for early pregnancy loss.** *Epidemiology.* 1990; **1**(5): 382–5.
PubMed Abstract

76. Sapra KJ, Buck Louis GM, Sundaram R, *et al.*: **Signs and symptoms associated with early pregnancy loss: findings from a population-based preconception cohort.** *Hum Reprod.* 2016; **31**(4): 887–96.
PubMed Abstract | Publisher Full Text | Free Full Text

77. Wilcox AJ, Baird DD, Weinberg CR: **Time of implantation of the conceptus and loss of pregnancy.** *N Engl J Med.* 1999; **340**(23): 1796–9.
PubMed Abstract | Publisher Full Text

78. Wilcox AJ, Dunson DB, Weinberg CR, *et al.*: **Likelihood of conception with a single act of intercourse: providing benchmark rates for assessment of post-coital contraceptives.** *Contraception.* 2001; **63**(4): 211–5.
PubMed Abstract | Publisher Full Text

79. Wilcox AJ, Weinberg CR, Baird DD: **Timing of sexual intercourse in relation to ovulation. Effects on the probability of conception, survival of the pregnancy, and sex of the baby.** *N Engl J Med.* 1995; **333**(23): 1517–21.
PubMed Abstract | Publisher Full Text

80. Wilcox AJ, Baird DD, Dunson D, *et al.*: **Natural limits of pregnancy testing in relation to the expected menstrual period.** *JAMA.* 2001; **286**(14): 1759–61.
PubMed Abstract | Publisher Full Text

81. Weinberg CR, Gladen BC, Wilcox AJ: **Models relating the timing of intercourse to the probability of conception and the sex of the baby.** *Biometrics.* 1994; **50**(2): 358–67.
PubMed Abstract | Publisher Full Text

82. Weinberg CR, Moledor E, Baird DD, *et al.*: **Is there a seasonal pattern in risk of early pregnancy loss?** *Epidemiology.* 1994; **5**(5): 484–9.
PubMed Abstract

83. Weinberg CR, Hertz-Picciotto I, Baird DD, *et al.*: **Efficiency and bias in studies of early pregnancy loss.** *Epidemiology.* 1992; **3**(1): 17–22.
PubMed Abstract | Publisher Full Text

84. Barrett JC, Marshall J: **The risk of conception on different days of the menstrual cycle.** *Popul Stud (Camb).* 1969; **23**(3): 455–61.
PubMed Abstract | Publisher Full Text

85. Schwartz D, Macdonald PD, Heuchel V: **Fecundability, coital frequency and the viability of Ova.** *Popul Stud (Camb).* 1980; **34**(2): 397–400.
PubMed Abstract | Publisher Full Text

86. Boklage CE: **The frequency and and survival probability of natural twin conceptions.** In: Keith LG, Papiernik E, Keith DM, Lukie B, editors. *Multiple Pregnancy: Epidemiology, Gestation and Perinatal Outcome.* New York: Parthenon Publishing Group; 1995; 41–50.
Reference Source

87. Hertwig O: **Beiträge zur Kenntniss der Bildung, Befruchtung und Theilung des thierischen Eies (Contributions to the knowledge of the formation, fertilization and division of the animal egg)**. *Morphol Jahrb.*1876; **1**: 347–434.

88. Steptoe PC, Edwards RG: **Birth after the reimplantation of a human embryo.** *Lancet.* 1978; **2**(8085): 366.
PubMed Abstract | Publisher Full Text

89. Clift D, Schuh M: **Restarting life: fertilization and the transition from meiosis to mitosis.** *Nat Rev Mol Cell Biol.* 2013; **14**(9): 549–62.
PubMed Abstract | Publisher Full Text | Free Full Text

90. Hertig AT, Rock J, Adams EC, *et al.*: **Thirty-four fertilized human ova, good, bad and indifferent, recovered from 210 women of known fertility; a study of biologic wastage in early human pregnancy.** *Pediatrics.* 1959; **23**(1 Part 2): 202–11.
PubMed Abstract

91. Hertig AT: **A fifteen-year search for first-stage human ova.** *JAMA.* 1989; **261**(3): 434–5.
PubMed Abstract | Publisher Full Text

92. Rock J, Hertig AT: **Some aspects of early human development.** *Am J Obstet Gynecol.* 1942; **44**(6): 973–83.
Publisher Full Text

93. Hertig AT, Rock J: **On a human blastula recovered from the uterine cavity 4 days after ovulation.** *Anat Rec.* 1946; **94**: 469.
PubMed Abstract

94. Hertig AT, Rock J: **A series of potentially abortive ova recovered from fertile women prior to the first missed menstrual period.** *Am J Obstet Gynecol.* 1949; **58**(5): 968–93, illust.
PubMed Abstract | Publisher Full Text

95. Hertig AT, Rock J: **Two human ova of the pre-villous stage, having a developmental age of about 8 and 9 days respectively.** *Contrib Embryol.* 1949; **33**(213–221): 169–86.
PubMed Abstract

96. Hertig AT, Adams EC, McKay DG, *et al.*: **A thirteen-day human ovum studied histochemically.** *Am J Obstet Gynecol.* 1958; **76**(5): 1025–40; discussion 40-3.
PubMed Abstract | Publisher Full Text

97. Hertig AT, Rock J: **Searching for early fertilized human ova.** *Gynecol Invest.* 1973; **4**(3): 121–39.
PubMed Abstract | Publisher Full Text

98. Barrett JC: **Fecundability and coital frequency.** *Popul Stud (Camb).* 1971; **25**(2): 309–13.
PubMed Abstract | Publisher Full Text

99. Clopper CJ, Pearson ES: **The use of confidence or fiducial limits illustrated in the case of the binomial.** *Biometrika.* 1934; **26**(4): 404–13.
Publisher Full Text

100. Agresti A, Coull BA: **Approximate is better than "exact" for interval estimation of binomial proportions.** *Am Stat.* 1998; **52**(2): 119–26.
Publisher Full Text

101. Edwards RG: **The Cleaving Embryo and the Blastocyst.** In: *Conception in the Human Female.* London: Academic Press; 1980; 668–766 at 47ff.

102. Buster JE, Bustillo M, Rodi IA, *et al.*: **Biologic and morphologic development of donated human ova recovered by nonsurgical uterine lavage.** *Am J Obstet Gynecol.* 1985; **153**(2): 211–7.
PubMed Abstract | Publisher Full Text

103. Formigli L, Formigli G, Roccio C: **Donation of fertilized uterine ova to infertile women.** *Fertil Steril.* 1987; **47**(1): 162–5.
PubMed Abstract | Publisher Full Text

104. Formigli L, Roccio C, Belotti G, *et al.*: **Non-surgical flushing of the uterus for pre-embryo recovery: possible clinical applications.** *Hum Reprod.* 1990; **5**(3): 329–35.
PubMed Abstract

105. Sauer MV, Bustillo M, Rodi IA, *et al.*: ***In-vivo* blastocyst production and ovum yield among fertile women.** *Hum Reprod.* 1987; **2**(8): 701–3.
PubMed Abstract

106. Richter MA, Haning RV Jr, Shapiro SS: **Artificial donor insemination: fresh versus frozen semen; the patient as her own control.** *Fertil Steril.* 1984; **41**(2): 277–80.
PubMed Abstract | Publisher Full Text

107. HFEA: **Fertility Treatment in 2013 - trends and figures.** *Human Fertilisation & Embryology Authority.* 2013.
Reference Source

108. Asher GW: **Reproductive cycles of deer.** *Anim Reprod Sci.* 2011; **124**(3–4): 170–5.
PubMed Abstract | Publisher Full Text

109. Bolton VN, Braude PR: **Development of the human preimplantation embryo *in vitro*.** *Curr Top Dev Biol.* 1987; **23**: 93–114.
PubMed Abstract

110. Jones HW Jr, Oehninger S, Bocca S, *et al.*: **Reproductive efficiency of human oocytes fertilized *in vitro*.** *Facts Views Vis Obgyn.* 2010; **2**(3): 169–71.
PubMed Abstract | Free Full Text

111. Bolton VN, Leary C, Harbottle S, *et al.*: **How should we choose the 'best' embryo? A commentary on behalf of the British Fertility Society and the Association of Clinical Embryologists.** *Hum Fertil (Camb).* 2015; **18**(3): 156–64.
PubMed Abstract | Publisher Full Text

112. Machtinger R, Racowsky C: **Morphological systems of human embryo assessment and clinical evidence.** *Reprod Biomed Online.* 2013; **26**(3): 210–21.
PubMed Abstract | Publisher Full Text

113. Niakan KK, Han J, Pedersen RA, *et al.*: **Human pre-implantation embryo development.** *Development.* 2012; **139**(5): 829–41.
PubMed Abstract | Publisher Full Text | Free Full Text

114. Koot YE, Teklenburg G, Salker MS, *et al.*: **Molecular aspects of implantation failure.** *Biochim Biophys Acta.* 2012; **1822**(12): 1943–50.
PubMed Abstract | Publisher Full Text

115. Daya S, Gunby J, Hughes EG, *et al.*: **Natural cycles for *in-vitro* fertilization: cost-effectiveness analysis and factors influencing outcome.** *Hum Reprod.* 1995; **10**(7): 1719–24.
PubMed Abstract

116. Zayed F, Lenton EA, Cooke ID: **Natural cycle *in-vitro* fertilization in couples with unexplained infertility: impact of various factors on outcome.** *Hum Reprod.* 1997; **12**(11): 2402–7.
PubMed Abstract | Publisher Full Text

117. Bassil S, Godin PA, Donnez J: **Outcome of *in-vitro* fertilization through natural cycles in poor responders.** *Hum Reprod.* 1999; **14**(5): 1262–5.
PubMed Abstract | Publisher Full Text

118. Roesner S, Pflaumer U, Germeyer A, *et al.*: **Natural cycle IVF: evaluation of 463 cycles and summary of the current literature.** *Arch Gynecol Obstet.* 2014; **289**(6): 1347–54.
PubMed Abstract | Publisher Full Text

119. Omland AK, Fedorcsák P, Storeng R, *et al.*: **Natural cycle IVF in unexplained, endometriosis-associated and tubal factor infertility.** *Hum Reprod.* 2001; **16**(12): 2587–92.
PubMed Abstract | Publisher Full Text

120. Janssens RM, Lambalk CB, Vermeiden JP, *et al.*: ***In-vitro* fertilization in a spontaneous cycle: easy, cheap and realistic.** *Hum Reprod.* 2000; **15**(2): 314–8.
PubMed Abstract | Publisher Full Text

121. Fahy UM, Cahill DJ, Wardle PG, *et al.*: ***In-vitro* fertilization in completely natural cycles.** *Hum Reprod.* 1995; **10**(3): 572–5.
PubMed Abstract

122. Braude PR, Johnson MH, Pickering SJ, *et al.*: **Mechanisms of Early Embryonic Loss *In Vivo* and *In Vitro*.** In: Chapman. M, Grudzinskas G, Chard T, editors. *The Embryo: Normal and Abnormal Development and Growth.* London: Springer-Verlag; 1991; 1–10.
Publisher Full Text

123. Dunson DB, Colombo B, Baird DD: **Changes with age in the level and duration of fertility in the menstrual cycle.** *Hum Reprod.* 2002; **17**(5): 1399–403.
PubMed Abstract | Publisher Full Text

124. Wilcox AJ, Weinberg CR, Baird DD: **Post-ovulatory ageing of the human oocyte and embryo failure.** *Hum Reprod.* 1998; **13**(2): 394–7.
PubMed Abstract | Publisher Full Text

125. Edwards RG: **Sexuality and Coitus.** In: *Conception in the Human Female.* London: Academic Press; 1980; 525–72 at 60ff.

126. Ioannidis JP: **Why most published research findings are false.** *PLoS Med.* 2005; **2**(8): e124.
PubMed Abstract | Publisher Full Text | Free Full Text

127. McCoy RC, Demko ZP, Ryan A, *et al.*: **Evidence of Selection against Complex Mitotic-Origin Aneuploidy during Preimplantation Development.** *PLoS Genet.* 2015; **11**(10): e1005601.
PubMed Abstract | Publisher Full Text | Free Full Text

128. Harris J: **Germline Modification and the Burden of Human Existence.** *Camb Q Healthc Ethics.* 2016; **25**(1): 6–18.
PubMed Abstract | Publisher Full Text

129. Saravelos SH, Regan L: **Early pregnancy failure after assisted reproductive technology.** *Pregnancy after Assisted Reproductive Technology.* 2012; 51–65.

130. Jones DG, Towns CR: **Navigating the quagmire: the regulation of human embryonic stem cell research.** *Hum Reprod.* 2006; **21**(5): 1113–6.
PubMed Abstract | Publisher Full Text

131. Dupont C, Froenicke L, Lyons LA, *et al.*: **Chromosomal instability in rhesus macaque preimplantation embryos.** *Fertil Steril.* 2009; **91**(4): 1230–7.
PubMed Abstract | Publisher Full Text | Free Full Text

132. Daughtry BL, Chavez SL: **Chromosomal instability in mammalian pre-implantation embryos: potential causes, detection methods, and clinical consequences.** *Cell Tissue Res.* 2016; **363**(1): 201–25.
PubMed Abstract | Publisher Full Text

133. Shorten PR, Peterson AJ, O'Connell AR, *et al.*: **A mathematical model of pregnancy recognition in mammals.** *J Theor Biol.* 2010; **266**(1): 62–9.
PubMed Abstract | Publisher Full Text

134. Potts M, Diggory P, Peel J: **Spontaneous Abortion.** *Abortion.* Cambridge: Cambridge University Press; 1977; 45–64.
Reference Source

135. Jarvis G: **Dataset 1 in: Early embryo mortality in natural human reproduction: What the data say.** *F1000Research.* 2016.
Data Source

136. Jarvis G: **Dataset 2 in: Early embryo mortality in natural human reproduction: What the data say.** *F1000Research.* 2016.
Data Source

137. Jarvis G: **Dataset 3 in: Early embryo mortality in natural human reproduction: What the data say.** *F1000Research.* 2016.
Data Source

138. Jarvis G: **Dataset 4 in: Early embryo mortality in natural human reproduction: What the data say.** *F1000Research.* 2016.
Data Source

Open Peer Review

Current Peer Review Status: ? ✔ ?

Version 1

Reviewer Report 23 May 2017

https://doi.org/10.5256/f1000research.9616.r22945

?

Steven H. Orzack
Fresh Pond Research Institute, Cambridge, MA, USA
James E. Zuckerman
Brigham and Women's Hospital, Harvard Medical School , Boston, MA, USA

Dr. Jarvis assesses the empirical support for the belief that there is a “great deal” of fetal wastage in humans. His conclusion is that there is less wastage than is often believed and that the percent loss between conception and birth is 40-60%. Resolution of this issue is important, as it has substantial implications for our understanding of early human development.

Dr. Jarvis describes present understanding as (p. 2):

Among reputable scientific publications, including medical and reproductive biology text books, scientific reviews and primary research articles, reported mortality estimates include: 30–70% before and during implantation; >50%, 73% and 80% before the 6th week; 75% before the 8th week; 70% in the first trimester; 40– 50% in the first 20 weeks; and 49%, >50%, 53%, 54%, 60%, >60%, 63%, 70%, 50–75%, 76%, 78%, 80–85%, >85%, and 90% total loss from fertilisation to term.

He states (p. 2) that four types of evidence underlie these claims:

1. A speculative hypothesis published in The Lancet.
2. Life tables of intra-uterine mortality.
3. Studies of early pregnancy by biochemical detection of hCG.
4. Anatomical studies of Dr Arthur Hertig and Dr John Rock.

On the basis of his review of this evidence, Dr. Jarvis concludes (p. 12) that “....10-40% is a plausible range for pre-implantation embryo loss and overall pregnancy loss from fertilization to birth is approximately 40-60%.”

This means that the best estimate of pre-birth mortality according to Dr. Jarvis is consistent with many previous estimates. In order to understand this consistency, it is useful to examine these types of evidence and what Dr. Jarvis makes of each. I discuss them in turn.

1. The Lancet article is Roberts & Lowe (1975). These authors concluded (p. 498) from their "speculative" analysis of the number of married women age 20-29 in England and Wales and of the number of live and dead births that 78% of conceptions are lost. In order to generate this estimate, the authors estimated the number of conceptions in any given year (based on the number of sexual encounters, probability of fertilization, etc.). Dr. Jarvis assesses the influence of changing the number of conceptions on the estimate of fetal wastage and shows (p. 3) that a low estimate of the number of conceptions results in an estimate of 22% conceptions lost and that a high estimate of the number of conceptions results in an estimate of 92% loss. He also generates a 95% confidence interval for the loss percentage of 37% - 90% by doing a simulation in which each value contributing to the number of conceptions is normally-distributed with a mean identical to Roberts and Lowe's value and a coefficient of variation of 20%. On this basis, he concludes about Roberts and Lowe's analysis that (p. 1) it "....has no quantitative value." and that (p. 4) it "....has no practical quantitative value".

Dr. Jarvis provides a useful sensitivity analysis of Roberts and Lowe's estimate, which should be taken seriously by those who may believe that their analysis is definitive (their paper has been cited more than 300 times, with many citations that point to the 78% estimate). That said, Dr. Jarvis' conclusion that Roberts and Lowe's analysis is quantitatively useless is itself incoherent. A number is a number and as a starting point, their estimate is useful although limited. If their analysis lacks "practical quantitative value" so too does the analysis of Dr. Jarvis. After all, there is no empirical basis for his assumptions about the statistical independence of the components contributing to his estimate of percentage or that these components are normally-distributed or that they have a coefficient of variation of 20%. It is not as though simply making arbitrary assumptions about the variability of parameters somehow means that an analysis is more quantitatively useful than one without such assumptions. The point is that both analyses have value. It is telling in this regard that their estimate is "close" to Dr. Jarvis' estimate. In fact, one could readily claim that Dr. Jarvis's analysis validates Roberts and Lowe's estimate in as much as their estimate is within the 95% confidence interval he generates.

By way of understanding Robert and Lowe's self-described "speculative" work, it is important to note it belongs to the voluminous "gray" literature relating to human pregnancy. This is the literature that is published without much review (if any) and without much requirement for rigor and data. To see this, one need go farther than this passage (p. 498):

Animal studies, which allow a more systematic investigation of [pregnancy loss], have shown detectable prenatal losses ranging from 15 to 60% in domestic cattle, sheep, and pigs and in wild forms such as stoats, rats, squirrels, and rabbits.

They cite Austin (1972) for this claim. He merely states (p. 134):

The data show that prenatal losses ranging between 15 and 60 per cent occur in cattle, sheep and pigs, as well as in wild forms such as stoats, rats, squirrels and rabbits.

No data are cited! In fact, Austin's gloss on the loss percentage for domesticated species is reasonably accurate (Casida, 1953; First & Eyestone, 1988; Lasley, 1957) although there are less data than one might imagine. It is of note that these species have been selected for offspring production and so how relevant these data are is not completely resolved. Perhaps fetal wastage in their wild relatives would be greater. My guess is that the data alluded to as being from "wild forms" are in papers such as those by Brambell (1942, 1948). That said, to my knowledge, it is not

clear that such studies reliably account for early gestational losses. More generally, there are few "wild forms" for which there are estimates.

The overall point is that Robert and Lowe's paper contains a disconnection between data and conclusions that would be sustained even if one read the cited source. Their paper is best viewed as a heuristic exercise. This is not a criticism. It is meant to underscore that Dr. Jarvis' conclusion that their paper is "useless" treats it as something that it isn't. We are ignorant of the training of Drs. Robert and Lowe but like many authors of the gray literature concerning pregnancy, they may have lacked rigorous training in research practice and data analysis. This is not inherently bad, as long as the nature of such publications is properly understood. As a community of scientists, we can make use of their insight into human pregnancy as long as its potential limitations are understood. We need all the help we can get!

2. The "life tables of intra-uterine mortality" are French & Bierman (1962) and Léridon (1977).The former study is an analysis of pregnancies in Kauai, Hawaii; the authors' conclusion was that approximately 24% of the pregnancies registered with an estimated gestational age of greater than four weeks would die. Léridon married this result with the data of Hertig, Rock, Adams, & Menkin (1959), which provide an estimate of wastage prior to four weeks, to infer that 63% of conceptions die before birth (Table 4.20, p. 81). Dr. Jarvis' cautions about the assumptions that underlie this estimate are reasonable. That said, it is important to note that the Léridon's chapter ("Intrauterine Mortality", pp. 48-81) is no casual exercise. It is the longest chapter in the book and an open-minded reader can see that Table 4.20 is based upon reasonable assumptions that Léridon clearly states do not have as much of a solid empirical basis as would be desired. Unfortunately, Dr. Jarvis' sole mentions of Léridon's caveats are a statement (p. 5) in which Léridon describes (p. 56) an interpolation he makes (in his analysis of French and Bierman's data) as "risky" and another in which his (Dr. Jarvis) reanalyses of the French and Bierman data (p. 5) "reinforce a concern highlighted by Léridon". To this extent, a reader of Dr. Jarvis' paper could easily come away with the mistaken belief that Léridon's analysis is superficial at best. As in the case of Roberts and Lowe's estimate, it is important to note that Léridon's estimate of conceptions lost of 63% is close to Dr. Jarvis' estimate of 40-60%.

3. "Studies of early pregnancy by biochemical detection of hCG." The modern pregnancy test is based upon an assay of human chorionic gonadotrophin (hCG), an oligosaccharide glycoprotein hormone produced by embryonic cells. An elevated level of hCG is detectable six to fourteen days post-conception (Nepomnaschy, Weinberg, Wilcox, & Baird, 2008; Wilcox, Baird, & Weinberg,1999). By this time, most embryos capable of implantation will have done so. Unfortunately, earlier pre-implantation detection of pregnancy based upon assay of the "Early Pregnancy Factor", a heat-shock protein expressed within 48 hours of conception, is not in widespread use (Clarke, 1997; Fan & Zheng, 1997; Morton, Rolfe, & Cavanagh, 1992; Rolfe, 1982; Shahani, Moniz, Chitlange, & Meherji, 1991; Shahani, Moniz, Gokral, & Meherji, 1995; Smart, Fraser, Roberts, Clancy, & Cripps, 1982). Dr. Jarvis correctly describes the pioneering hCG results of Wilcox *et al.* (1988) and others (as summarized in Table 3), which indicate that the percentage loss of conceptions after hCG detection is between approximately 20 and 60%, with many estimates between 30 and 40%; Dr. Jarvis concludes (p. 6) that this percentage loss is approximately 33%.

Dr. Jarvis goes on to estimate that the "...loss from fertilization to birth [is] 40- 60%"; this is based on the combination of three estimates based on hCG assay of percentage loss from conception to birth (35.7%: Wang *et al.*, 2003; 31.3%: 31.3%: Wilcox *et al.*, 1988; 31.3% Zinaman, Clegg, Brown, O'Connor, & Selevan, 1996) and his estimate (pp. 7-8) that the efficiency of implantation of embryos "...may be up to 90% efficient...." in order. He concludes that higher estimates of loss from fertilization to birth from the literature are "excessive".

Dr. Jarvis' estimate is likely an underestimate. There is strong circumstantial evidence that many more than 10% of embryos do not successfully implant, as discussed below. The implication of this is that Dr. Jarvis' estimate and the previous estimates are consistent. It is also worth noting that Dr. Jarvis uses an arbitrary estimate for implantation rate, even though he judges other analyses to be useless because they contain an arbitrary parameter estimate.

Dr. Jarvis goes on to criticize Boklage (1990) who estimated the percentage of unsuccessful conceptions based on an analysis of hCG data (see his Figure 2, p. 84). Dr. Jarvis is right to raise concerns (p. 8) that Boklage's analysis is less definitive than desired. In particular, he states (p. 8) that Boklage's assumption that the 21-day survival rate of conceptions is 28.7% is based upon a misinterpretation of a previous study. That said, Dr. Jarvis makes an unsubstantiated conclusion (p. 8) that "...quantitative conclusions from [Boklage's] analysis in relation to the survival of naturally conceived human embryos are of doubtful validity". This may be true, but this remains to be seen given the lack of any demonstration of the sensitivity of Boklage's quantitative conclusions to changes in the underlying assumptions. Boklage's analysis needs more careful scrutiny than given by Dr. Jarvis. For example, Boklage presents a formula for the percentage loss of conceptions as a function of time (p. 84). Are the coefficients estimated via a standard statistical approach such as maximum likelihood estimation and chosen via a likelihood ratio test or via comparison of AIC values associated with competing models? This is not clear. As such, it is unclear as to what to make of the predictions even putting aside Dr. Jarvis' concerns about the biological validity of some of the underlying data. The equation appears to be based upon the assumption that a cohort of embryos is an admixture of those that are likely to die before six weeks and those that will survive longer. The basis for this assumption is unclear. The lack of transparency of Boklage's equation is underscored by the fact that Dr. Jarvis does not mention that it predicts 75.8 percent fetal wastage between conception and full-term birth (270 days). As above, this estimate is rightly or wrongly consistent with most previous estimates.

4. The "anatomical studies of Dr Arthur Hertig and Dr John Rock" are investigations of conceptions recovered from uteri obtained via gynecologic surgery. Their results are summarized in Hertig *et al.* (1959); Hertig & Rock, (1973); Hertig, (1967). As described by Dr. Jarvis (p. 9), Hertig *et al.*'s conclusion is that 50% of embryos will die within two weeks after conception.

Dr. Jarvis' is correct to point out concerns about their conclusion, although we believe that it has been well recognized that it is "impressionistic" as opposed to something that has a solid quantitative underpinning. Of course, as noted by Dr. Jarvis, their work remains important.

Dr. Jarvis makes some assertions about Hertig *et al.*'s work that seem mainly intended to accentuate doubts about it as opposed to placing it in proper context. He notes correctly (p. 9) that the sample is cross-sectional and not longitudinal. Given the nature of this study, this was

unavoidable. Dr. Jarvis notes there are some unresolved discrepancies among age-specific detection rates for embryos and also between the estimated implantation rate and the rate inferred from other studies. These are worth mentioning but the implications of these discrepancies remain ambiguous in the absence of a quantitative analysis that accounts for sampling variation.

Similarly un-useful is Dr. Jarvis' statement (p. 9) that "Despite having proven fertility, these women presented with gynaecological problems, suggesting suboptimal reproductive function." There is a wide range of "gynaecological problems" and an unanchored assertion that such a broad category might result in "sub-optimal reproductive function" means nothing in the absence of evidence that whatever problems were present had some influence on embryonic viability. In an effort to "estimate the precision" of the various proportions presented by Hertig *et al.* (e.g., the survival rate to implantation), Dr. Jarvis generated 500 so called "bootstrap" samples from the original data consisting of 107 cases. These samples arise from sampling with replacement of the original data (e.g., see Efron & Tibshirani, 1986; Efron, 1987). Such an investigation is worthwhile, although a bootstrap analysis is not a "cure" for small sample size. In any case, Dr. Jarvis' analyses of the bootstrap results are incorrect. He describes (p. 10) "95% CIs" for various proportions that are outside of the range of 0-100%. For example, the confidence interval (p. 10) he provides for pre-implantation embryo survival probability is 27-128%. Such an interval cannot be generated by a correct bootstrap analysis. There are various ways to calculate a bootstrap confidence interval (Efron & Tibshirani, 1986). The simplest, known as the "percentile method", generates a 95% bootstrap confidence interval for a proportion directly from the range of proportions associated with the central 95% of the bootstrap estimates. Accordingly, the confidence interval must be between 0 and 100% because each of the bootstrap samples must generate a proportion between 0 and 100%. Dr. Jarvis' mistake appears to be that he estimated an average proportion and its variance from the ensemble of bootstrap estimates and then calculated the confidence interval using standard formulae (p. 10). The purpose of bootstrap estimation is to avoid such calculations, which can generate inaccurate confidence intervals. Although some of the bootstrap confidence intervals provided by Dr. Jarvis do not fall below 0% or surpass 100%, we guess that all of them are incorrectly calculated. Unfortunately, the incorrect confidence intervals are described by Dr. Jarvis (p. 12) as "mathematically and biologically implausible" and taken to "....betray the quantitative weaknesses in [Hertig *et al.*'s] data and analysis." Indeed, they are "mathematically and biologically implausible" but the reason is that they were not correctly calculated. Whatever bearing a bootstrap analysis has on our understanding of the "precision" of Hertig *et al.*'s data and analyses remains to be seen.

Dr. Jarvis' central argument is that there is more ambiguity associated with estimates of fetal wastage in humans and that this ambiguity is not widely understood. Many of his concerns should be taken seriously. Nonetheless, his analysis is undermined by errors of analysis and overstatement. In the end, his estimate of fetal wastage from conception to birth is consistent with many of the previous estimates.

Dr. Jarvis' analysis is also undermined by an incorrect dismissal of data from embryos created via assisted reproductive technology (ART), which he refers to as *in vitro* fertilization (IVF). On page 11, he alludes to "...sub-optimal conditions for embryo culture..." and implies that somehow ART embryos are "different" in undefined ways from naturally-conceived embryos that negate their potential use in regard to estimating fetal wastage. This is an exercise in rhetoric, not a scientific

argument. It is true that ART embryos are different from natural embryos in ways that could influence an estimate of fetal wastage. However, it is essential to note that they constitute the best available sample for insight into the "black box" of early pregnancy, despite the possible biases they may have that could distort our view into the black box. To this extent, it is best to assess what information they can provide about fetal wastage, rather than provide tenuous or irrelevant reasons as to why they are not useful.

Dr. Jarvis mistakenly assumes (p. 11) that only ART embryos transferred into mothers would provide information about fetal wastage. In fact, as Dr. Jarvis notes, there are a number of reasons why transferred embryos are not representative of all embryos (e.g., conscious or unconscious quality biases, sex selection) and accordingly, this kind of sample could be misleading. That said, studies of such samples suggest that at least some aspects of their biology are identical to that of naturally-conceived embryos. For example, the sex ratio at birth for ART embryos is statistically identical with that of natural conceptions (Orzack *et al.*, 2015).

More importantly, the entire ensemble of ART embryos (untransferred and transferred) provides information about fetal wastage. Almost all ART embryos undergoe testing for chromosomal abnormalities, such as aneuploidy. The consequences of aneuploidy are well-known – it results in almost certain death before birth. This is consistent with the fact that many spontaneous abortions are karyotypically abnormal (Boué, Boué, & Lazar, 1967, 1975; Jauniaux & Burton, 2005). To this extent, the frequency of such abnormalities provides strong circumstantial evidence as to the amount of fetal wastage. Orzack *et al.* (2015) investigated a sample of ART embryos whose karyotypes were assayed via FISH or CGH and reported that 84,881 out of 139,704 embryos contained at least one aneuploid chromosome. The implied percentage of fetal wastage (60.8%) is remarkably consistent with the central tendency of the many reports that Dr. Jarvis dismisses as unreliable, as well as with his own estimate. As noted, we need to be cautious about inferences from this sample but not avoid making them. There is no compelling reason to think that "suboptimal" conditions for embryo culture (if any) cause many chromosomal abnormalities, most of which very likely arise during meiosis (e.g., Hassold & Hunt, 2001; Hunt & Hassold, 2007; Jones, 2008; Nagaoka, Hassold, & Hunt, 2012). What deserves scrutiny are whether the frequency of chromosomal abnormalities is elevated by techniques for collecting eggs and/or because women providing them for use in ART are unrepresentative of all reproductive women. There are limited data that unstimulated and stimulated oocytes have similar frequencies of abnormality (Labarta *et al.*, 2010). Of course, women using ART are often older than many typical mothers. However, a high frequency of karyotypic abnormality is also observed among oocytes from young women (Baart *et al.*, 2006; Munné *et al.*, 2006). These concerns should continue to be investigated but they in no way imply that ART embryos cannot provide useful insights about early human development and fetal wastage, especially given the current lack and very likely continuing lack of a large sample of naturally-conceived human embryos.

We see then a web of circumstantial evidence implying that there is a substantial amount of fetal wastage in humans. This insight arises from imperfect types of knowledge (as documented by Dr. Jarvis) but nonetheless, there is a signal consistent with the claim that approximately half or more of conceptions fail. More needs to be done to improve our understanding.

The study of fetal wastage shares with the study of the human sex ratio during pregnancy the fact that many different kinds of scientists are involved and so, the associated balkanization has reduced the accountability that arises from a shared disciplinary perspective about the standards

for the interpretation of data (Orzack, 2016; Orzack *et al.*, 2015). One cause and consequence of this division is the gray
literature mentioned above.

What contributes to the continuing "life" of the gray literature? Science abhors a vacuum and claims about high fetal wastage in humans have been repeated often in a way that the connection with assumptions and data have gotten obscured or lost. Some claims date well before there was any means by which early mortality could be assessed (Mall, 1917; Meyer, 1920; Pearson, 1897). Pearson clearly acknowledged the lack of direct evidence but such caveats get lost especially in medicine in which attention to standards of evidence, recognition of the assumptions needed to connect data with conclusions, and awareness of needed statistical techniques have been less as compared to in biological research. These deficiencies have diminished as medical training has incorporated more scientific training but have not disappeared. Nonetheless, during medical training the "inhalation" of facts is important. It is one reason as to why many believe that fetal wastage is high, despite having little or no familiarity with the available data along with the ins and outs of their analysis and interpretation.

In this context, care is needed when assessing the nature of claims about fetal wastage. This can be illustrated by considering Dr. Jarvis' claim (p. 8):

....it is clear that estimates for total embryonic loss of 90% (Opitz, 2002), 85% (Braude & Johnson, 1990), 83% (Harris, 2003), 80– 85% (Johnson & Everitt, 2000; Vitzthum, Spielvogel, Thornburg, & West, 2006), 78% (Roberts & Lowe, 1975), 76% (Boklage, 1990; Drife, 1983) and 70% (Chard, 1991; Ford & Schust, 2009; Loke & King, 1995; Macklon, Geraedts, & Fauser, 2002; McCoy et al., 2015) are excessive.

(We have replaced number citations with author citations). Several of these claims are in medical textbooks and are akin to newspaper articles, i.e., they are reports on prior research as opposed to being independent estimates. Even then the nature of the evidence can go unmentioned. For example, in their text book Johnson & Everitt (2000) include no evidence or citations in which to find evidence
underlying their estimate. Of the claims in the primary literature, we again see a lack of independent evidence in as much as someone else's estimate is reported. For example, Chard (1991); Drife (1983); Vitzthum *et al.* (2006) merely present Roberts & Lowe's (1975) estimate. A few claims present their own evidence. For example, Harris (2003) contains this passage (p. 362):

We now know that for every successful pregnancy that results in a live birth many, perhaps as many as five early embryos will be lost or will "miscarry"....

and accompanying footnote (p. 371):

Robert Winston gave the figure of five embryos for every live birth some years ago in a personal communication. Anecdotal evidence to me from a number of sources confirms this high figure, but the literature is rather more conservative, making more probable a figure of three embryos lost for every live birth. See: Boklage CE. Survival probability of human conceptions from fertilization to term. International Journal of Fertility 1990;35(2)75–94. See also: Leridon H. Human Fertility: The Basic Components. Chicago: University of Chicago Press; 1977. Again, in a recent personal communication, Henri Leridon confirmed that a figure of three lost embryos for every live birth is a reasonable conservative figure.

This is clearly a heuristic estimate! The point is that there is less of a monolithic ensemble of flawed estimates that need to be debunked than one might imagine given Dr. Jarvis' passage. In any case, there is nothing inherently problematic about the citations just described. Indeed, it would be preferable if attributions were better and speculation was better highlighted as such. Nonetheless, such estimates should be used with caution but not discarded, given the substantial difficulties associated with the estimation of fetal wastage in humans.

An ideal future investigation of fetal wastage is easy to imagine: daily assessment of EPF and hCG for a cohort of women attempting to get pregnant. Easier said than done! Consider what such a study would require: a reliable assay for EPF, the enrollment of thousands of women, collection of and accurate assessment of thousands of samples, and more. Perhaps these technical and logistical barriers can be overcome soon. In the meantime, we can recognize that there is strong circumstantial evidence that human fetal wastage is likely between 50 and 75%. At the same time, we can recognize along with Dr. Jarvis that this conclusion lacks definitive proof and that additional investigations and scrutiny are needed.

References

1. Austin C. R.: Pregnancy losses and birth defects.*Reproduction in Mammals 2: Embryonic and Fetal Development*. 1972. 134-153
2. Baart EB, Martini E, van den Berg I, Macklon NS, et al.: Preimplantation genetic screening reveals a high incidence of aneuploidy and mosaicism in embryos from young women undergoing IVF.*Hum Reprod*. 2006; **21** (1): 223-33 PubMed Abstract | Publisher Full Text
3. Boklage CE: Survival probability of human conceptions from fertilization to term.*Int J Fertil*. **35** (2): 75, 79-80, 81 PubMed Abstract
4. Boué JG, Boué A, Lazar P: [Chromosome aberrations in abortions].*Ann Genet*. 1967; **10** (4): 179-87 PubMed Abstract
5. Boué J, Bou A, Lazar P: Retrospective and prospective epidemiological studies of 1500 karyotyped spontaneous human abortions.*Teratology*. 1975; **12** (1): 11-26 PubMed Abstract | Publisher Full Text
6. Brambell F.W.R: Intra-Uterine Mortality of the Wild Rabbit, Oryctolagus cuniculus (L.). *Proceedings of the Royal Society of London. Series B - Biological Sciences*. 1942; **130** (861): 462-479 Reference Source
7. BRAMBELL FW: Prenatal mortality in mammals.*Biol Rev Camb Philos Soc*. 1948; **23** (4): 370-407 PubMed Abstract
8. Braude P. R, Johnson M. H.: The embryo in contemporary medical science. *The Human Embryo: Aristotle and the Arabic and European Traditions*. 1990. 208-221
9. Casida L. E.: Fertilization failure and embryonic death in domestic animals.*Pregnancy Wastage. Thomas, Springfield, Ill*. 1953. 27-37
10. Chard T: Frequency of implantation and early pregnancy loss in natural cycles.*Baillieres Clin Obstet Gynaecol*. 1991; **5** (1): 179-89 PubMed Abstract
11. Clarke FM: Controversies in assisted reproduction and genetics. Does. *J Assist Reprod Genet*. 1997; **14** (9): 489-91 PubMed Abstract
12. Drife J. O.: What proportion of pregnancies are spontaneously aborted?. *British Medical Journal, 286*. 1983. 294
13. Efron B.: Better bootstrap confidence intervals. *Journal of the American statistical Association*. 1987; **82** (397): 171-185 Reference Source
14. Efron B., Tibshirani R. J.: Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy.*Statistical Science*. 1986; **1** (1): 54-75 Reference Source

15. Fan XG, Zheng ZQ: A study of early pregnancy factor activity in preimplantation.*Am J Reprod Immunol*. 1997; **37** (5): 359-64 PubMed Abstract
16. First N. L., Eyestone W. H.: Reproductive efficiency in domestic animals. *Annals of the New York Academy of Science*. 1988; **541** (1): 697-705 Reference Source
17. Ford HB, Schust DJ: Recurrent pregnancy loss: etiology, diagnosis, and therapy.*Rev Obstet Gynecol*. 2009; **2** (2): 76-83 PubMed Abstract
18. FRENCH FE, BIERMAN JM: Probabilities of fetal mortality.*Public Health Rep*. 1962; **77**: 835-47 PubMed Abstract
19. Harris J: Stem cells, sex, and procreation.*Camb Q Healthc Ethics*. 2003; **12** (4): 353-71 PubMed Abstract
20. Hassold T, Hunt P: To err (meiotically) is human: the genesis of human aneuploidy.*Nat Rev Genet*. 2001; **2** (4): 280-91 PubMed Abstract | Publisher Full Text
21. Hertig A. T.: The overall problem in man. *In Comparative Aspects ofReproductive Failure Springer Berlin Heidelberg*. 1967. 11-41 Reference Source
22. Hertig AT, Rock J: Searching for early fertilized human ova.*Gynecol Invest*. 1973; **4** (3): 121-39 PubMed Abstract
23. HERTIG AT, ROCK J, ADAMS EC, MENKIN MC: Thirty-four fertilized human ova, good, bad and indifferent, recovered from 210 women of known fertility; a study of biologic wastage in early human pregnancy.*Pediatrics*. 1959; **23** (1 Part 2): 202-11 PubMed Abstract
24. Hunt PA, Hassold TJ: Human female meiosis: what makes a good egg go bad?. *Trends Genet*. 2008; **24** (2): 86-93 PubMed Abstract | Publisher Full Text
25. Jauniaux E, Burton GJ: Pathophysiology of histological changes in early pregnancy loss.*Placenta* . **26** (2-3): 114-23 PubMed Abstract | Publisher Full Text
26. Johnson M. H., Everitt B. J.*Essential Reproduction (5th ed.)*. 2000.
27. Jones KT: Meiosis in oocytes: predisposition to aneuploidy and its increased incidence with age. *Hum Reprod Update*. **14** (2): 143-58 PubMed Abstract | Publisher Full Text
28. Labarta E., Bosch E., Alama P., Rubio C., et al.: Ovarian stimulation does not increase embryo aneuploidy rates in young normo-ovulatory women.*Human Reproduction*. 2010; **25**: I62-I63
29. Lasley E. L.: Ovulation, prenatal mortality and litter size in swine. *Journal of Animal Science*. 1957; **2** (16): 335-340 Reference Source
30. Léridon H: Human fertility: The Basic Components.*Chicago: University of Chicago Press.*1977.
31. Loke Y. W., King A.: Human Implantation: Cell Biology and immunology.*Cambridge:Cambridge University Press.*1995.
32. Macklon NS, Geraedts JP, Fauser BC: Conception to ongoing pregnancy: the 'black box' of early pregnancy loss.*Hum Reprod Update*. **8** (4): 333-43 PubMed Abstract
33. Mall F. P.: Note on abortions with letters from the Health Commissioner of Baltimore and from the Chief of the Bureau of Vital Statistics of Maryland regarding registration andshipment of embryos to the Carnegie Laboratory of Embryology at the Johns Hopkins Medical. Baltimore.1917.
34. McCoy RC, Demko Z, Ryan A, Banjevic M, et al.: Common variants spanning PLK4 are associated with mitotic-origin aneuploidy in human embryos.*Science*. 2015; **348** (6231): 235-8 PubMed Abstract | Publisher Full Text
35. Meyer AW: THE PRACTITIONER'S CONTRIBUTION TO EMBRYOLOGY.*Cal State J Med*. 1920; **18** (2): 44-6 PubMed Abstract
36. Morton H., Rolfe B. E., Cavanagh A. C.: Early pregnancy factor. *Seminars in Reproductive Endocrinology*. 1992; **10** (2): 72-82
37. Munné S, Ary J, Zouves C, Escudero T, et al.: Wide range of chromosome abnormalities in the embryos of young egg donors.*Reprod Biomed Online*. 2006; **12** (3): 340-6 PubMed Abstract
38. Nagaoka SI, Hassold TJ, Hunt PA: Human aneuploidy: mechanisms and new insights into an

age-old problem.*Nat Rev Genet*. 2012; **13** (7): 493-504 PubMed Abstract | Publisher Full Text
39. Nepomnaschy PA, Weinberg CR, Wilcox AJ, Baird DD: Urinary hCG patterns during the week following implantation.*Hum Reprod*. 2008; **23** (2): 271-7 PubMed Abstract | Publisher Full Text
40. Opitz J. M.: Human development - the long and the short of it. In E. J. Furton. 2002.
41. Orzack S. H.: Old and new ideas about the human sex ratio. *Significance*. 2016; **13** (1): 24-27 Publisher Full Text
42. Orzack SH, Stubblefield JW, Akmaev VR, Colls P, et al.: The human sex ratio from conception to birth.*Proc Natl Acad Sci U S A*. 2015; **112** (16): E2102-11 PubMed Abstract | Publisher Full Text
43. Pearson K.: The Chances of Death and Other Studies in Evolution. *CUP Archive*. 1897; **1**.
44. Roberts C. J., Lowe C. R.: Where have all the conceptions gone?. *Problems of Birth Defects*. 1975. 148-150
45. Rolfe BE: Detection of fetal wastage.*Fertil Steril*. 1982; **37** (5): 655-60 PubMed Abstract
46. Shahani SK, Moniz C, Chitlange S, Meherji P: Early pregnancy factor (EPF) as a marker for the diagnosis of subclinical embryonic loss.*Exp Clin Endocrinol*. 1992; **99** (3): 123-8 PubMed Abstract | Publisher Full Text
47. Shahani SK, Moniz CL, Gokral JS, Meherji PK: Early pregnancy factor (EPF) as a marker for detecting subclinical embryonic loss in clomiphene citrate-treated women.*Am J Reprod Immunol*. 1995; **33** (5): 350-3 PubMed Abstract
48. Smart YC, Fraser IS, Roberts TK, Clancy RL, et al.: Fertilization and early pregnancy loss in healthy women attempting conception.*Clin Reprod Fertil*. 1982; **1** (3): 177-84 PubMed Abstract
49. Vitzthum VJ, Spielvogel H, Thornburg J, West B: A prospective study of early pregnancy loss in humans.*Fertil Steril*. 2006; **86** (2): 373-9 PubMed Abstract | Publisher Full Text
50. Wang X, Chen C, Wang L, Chen D, et al.: Conception, early pregnancy loss, and time to clinical pregnancy: a population-based prospective study.*Fertil Steril*. 2003; **79** (3): 577-84 PubMed Abstract
51. Wilcox AJ, Baird DD, Weinberg CR: Time of implantation of the conceptus and loss of pregnancy.*N Engl J Med*. 1999; **340** (23): 1796-9 PubMed Abstract | Publisher Full Text
52. Wilcox AJ, Weinberg CR, O'Connor JF, Baird DD, et al.: Incidence of early loss of pregnancy.*N Engl J Med*. 1988; **319** (4): 189-94 PubMed Abstract | Publisher Full Text
53. Zinaman MJ, Clegg ED, Brown CC, O'Connor J, et al.: Estimates of human fertility and pregnancy loss.*Fertil Steril*. 1996; **65** (3): 503-9 PubMed Abstract

Competing Interests: No competing interests were disclosed.

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Author Response 31 May 2017

Gavin Jarvis, University of Sunderland, UK

I would like to thank Professor Orzack and Professor Zuckerman for their extensive review. I have amended the article in light of their remarks, and hope they consider it improved as a result.

Introduction

The purpose of my article is to evaluate available data that contribute to our quantitative understanding of natural human embryo mortality. The body of relevant data is small, as noted by the reviewers, although I have attempted to identify all of it. I deliberately avoided IVF/ART data since there is so much, and it is not obvious how such data illuminate natural circumstances (I comment further on this below). My comments on IVF/ART data are therefore confined to the Discussion.

Orzack & Zuckerman repeatedly refer to my estimates of 10-40% preimplantation loss and 40-60% total embryo loss, which are important benchmarks for my article. They are critical of these, although they do not seem to appreciate where they come from. Contrary to what they imply ("On the basis of his review of this evidence..."), they do not arise from analyses described in this article. Rather, they are from an analysis described in a previous article in *F1000Research*[1]. I have amended the article to clarify this point. Concerns with the validity of these estimates should focus on that analysis, which is not listed among their 53 references.

In their review, the reviewers are ambiguous (one might say 'gray') in their use of quotation marks and appear to ascribe to me things I did not write. For example, I do not use the phrase "great deal". Thus, for the sake of clarity, and to separate literary emphasis from quotation, I will follow the convention employed by GEM Anscombe, who coined a useful phrase[2], to distinguish between 'scare quotes' and "quotations".

I address points raised in the review, approximately in the order in which they appear.

1. Roberts & Lowe

Orzack & Zuckerman state that I calculate 95% confidence intervals. This is incorrect. The range [37-90%] is not a confidence interval, I do not refer to it as such, and nor can it be, since there are no data. As described in the article, it is the range within which 95% of simulated estimates fall, based on Roberts & Lowe's speculative values and other assumptions.

The reviewers suggest that my analysis lacks "practical quantitative value". I agree. This is the point and I am glad they have recognised it, if not entirely appreciated its significance. My analysis has "no practical quantitative value" *for estimating the number of conceptions that are lost*. As I explicitly point out, the sole purpose of the sensitivity analyses is to show that modest changes in the speculative estimates used by Roberts & Lowe may result in any biologically plausible value for embryo loss.

That my simulated estimate of 76.5% is close to Roberts & Lowe's 78% is not telling since it uses their original speculative values. On the contrary, it would be telling (of something) if they were not close. I simply added variance to the speculative values. I comment on the nature of this variance/covariance in my response to Reviewer 2. Thus, my analysis does not validate Roberts & Lowe, it exposes its quantitative futility.

Gray Literature is "documentary material which is not commercially published or publicly available, such as technical reports or internal business documents."[3] *The Lancet* is not 'Gray

Literature'. I comment further on this below. Contrary to the reviewers' suggestion, we are not completely "ignorant of the training of Drs. Roberts & Lowe" or unaware of their experience in "research practice and data analysis". Charles Ronald Lowe was the more senior of the two. He was 63 years old and Professor of Social and Occupational Medicine at the University of Wales College of Medicine when *The Lancet* article was published. He "contributed much to the growth of academic public health and the teaching of epidemiology and statistics."[4]

I do not describe their work as "useless" – if intended as a quote, then it is a misquote. I describe it as having "no practical quantitative value". These are carefully chosen words. (I have edited the equivalent phrase in the Abstract to match the full text.) The critique offered by the reviewers and their description of the paper as heuristic support this view. Nevertheless, I have added a statement that, as a model for highlighting factors that influence fecundity, the Roberts & Lowe analysis has some value.

In all fairness, on four separate occasions, I describe the analysis of Roberts & Lowe as a "hypothesis", i.e., the banner under which it was originally published in *The Lancet*. Indeed, they describe their arithmetic as "speculative"; however, they also describe their estimate as "conservative", implying that the true result may be even higher than 78%. My critique would be less germane had their hypothesis not been cited so widely ("more than 300 times", as helpfully pointed out by the reviewers). I suggest that it is not I, but those who enthusiastically cite it[5] who treat it as "something that it isn't".

2. Life Tables of Intrauterine Mortality

I do not consider Leridon's chapter[6] a "casual exercise" or "superficial". On the contrary, it is a well-reasoned attempt to answer a challenging biological question. I have included a tribute in my article to Leridon's review. I hope this prevents readers from gaining such false impressions.

I agree with the reviewers that Leridon's 63% is close to my 40-60%. However, Roberts & Lowe's 78% is not, as they imply.

A critique of Leridon's life-table is not a critique of Leridon at all, but of French & Bierman[7] and Hertig[8]. I discuss briefly why French & Bierman may be an overestimate and, in detail, how Hertig's analysis is flawed. Leridon's account has been widely cited, especially by those describing embryo loss at the earliest stage. I hope readers will find it useful to know how Leridon's values are derived.

3. hCG studies of early pregnancy loss

The Edmonds (1982) estimate of approximately 60% loss[9] is the highest I report and, for reasons discussed in the article and mentioned by others[10], is likely to be an over-estimate. Nevertheless, years after the more credible Wilcox (1988)[10] study was published, Edmonds is still widely cited to justify high levels of embryo wastage. For example, Hyde & Schust

(2015)[11] cite both Edmonds and Wilcox to support their claim that "Approximately 70% of human conceptions fail to achieve viability, with almost 50% of all pregnancies ending in miscarriage before the clinical recognition of a missed period..." By showing Edmonds' results in context, I hope this kind of overstatement can be avoided.

My conclusion of one third loss is based on the average of the eight listed studies from Wilcox to the present day (unweighted average = 31.9%). I have edited the paper to make this clear. I also discuss why the estimates prior to Wilcox are less reliable and cite several studies that make similar observations.

As already noted, my 40-60% estimate is from a previous analysis[1] and is not a combination of the values (31.3; 35.7; 31.3) highlighted by the reviewers. My rationale for using a 90% implantation (and fertilisation) efficiency is found in that analysis[1].

My conclusion regarding the validity of Boklage's analysis of embryo mortality[12] is not "unsubstantiated". Indeed, the reviewers mention a key point of substance: namely, that Boklage's value of 28.7% misinterprets the biology. Boklage uses this as measure of embryo mortality, whereas it is a fecundability. If fecundabilities are analysed as embryo mortalities, surely this casts doubt on the validity of conclusions regarding embryo mortality.

I cannot comment on Boklage's statistical methodology (i.e., use of MLE, LRTs or AIC values) since he reports no such detail. However, I thank the reviewers for highlighting the lack of clarity in Boklage's analysis.

Contrary to the claim of the reviewers, I refer to Boklage's estimate of 76% loss from conception (fertilisation) to birth on three occasions. This 76% estimate is consistent with Roberts & Lowe's value. It is somewhat higher than Leridon's (whose life table is inexplicably omitted from the Boklage analysis). It is clearly not consistent with my 40-60% estimate[1].

4. Hertig's data and analysis

Regarding Hertig's conclusion, Orzack & Zuckerman "believe that it has been well recognized that it is 'impressionistic' as opposed to something that has a solid quantitative underpinning". I agree that Hertig's conclusion does not have a "solid quantitative underpinning"; however, it is precisely the quantitative underpinning of Leridon's life table and other claims about early natural embryo mortality. This is a key point of my article. It is not clear what the reviewers mean by 'impressionistic'[13]: some authors seem to offer an 'unimpressionistic' account of Hertig. For example, in the widely-cited 'Black Box' review[14], Macklon *et al*. write regarding Hertig's study: "...the high rate of early pregnancy loss before the time of the first missed period was thus clearly demonstrated..." Other less widely-cited articles[15] do address the design and analytical shortcomings.

Pointing out shortcomings in studies is what scientists (and reviewers) are meant to do. Thus, I agree with the reviewers that they are "worth mentioning". Furthermore, by pointing out that Hertig's subjects were of proven fertility, had gynaecological problems and may have had suboptimal reproductive function, I am placing Hertig's study "in proper context".

This is not "un-useful". Nevertheless, I have edited this section, to accommodate these reviewers' scepticism with the more positive view of others[16]. I hope I have struck an acceptable balance.

Orzack & Zuckerman appear to have concerns with well-established statistical techniques, referring to my "so-called 'bootstrap' samples". I agree that bootstrapping is "not a 'cure' for small sample size", but I do not claim that it is. Bootstrapping can provide estimates of precision when it is not possible to calculate these analytically. As with all analyses, outputs require appropriate interpretation.

The reviewers state that the "analyses of the bootstrap results are incorrect" because some of the confidence intervals lie outside the range 0-100%. I am aware that this is impossible (for a probability) as I explicitly point out. Such outputs do indicate a serious flaw in the analysis, which is as follows: Hertig ignores 47 of his 107 cases. These cases are included in my bootstrap. The reader may consider whether ignoring 44% of the data is reasonable and the extent to which by doing so Hertig has generated biased estimates of the probabilities he calculates. Kline *et al*. (1989) make a similar point: "The missing data are sufficient to engender an entirely different result"[15]. The bootstrap therefore illustrates the extent to which Hertig's estimates are biased by ignoring his own data. There are other reasons to doubt the precision of his conclusions and the representative nature of the subset of data upon which he relies so heavily – these are described in the article.

The bootstrap pseudo-datasets are available for scrutiny (Dataset 4). Thus, if there are any flaws in my reasoning or bootstrap, the reviewers may point these out. I used the percentile method (to which they refer) to calculate the 95% CIs and I have edited the text to clarify this. I do not believe there are any flaws in my bootstrap.

IVF/ART data

There is a wealth of data from IVF/ART studies and I have only mentioned a tiny proportion of this. Orzack & Zuckerman and a previous reviewer[17] suggest that such data could contribute to a quantitative understanding of the *in vivo* situation. In the broadest sense, this is of course true. However, there are difficulties in extrapolating from *in vitro* to *in vivo* circumstances. I am not alone in pointing this out[14], and I have illustrated some of these difficulties in the Discussion.

My description of "sub-optimal conditions for embryo culture" is drawn from two papers:

1. Bolton & Braude (1987)[19]: "Optimal culture conditions for human embryos have yet to be defined" and "suboptimal culture conditions are undoubtedly responsible for a proportion of this embryonic failure".
2. Bolton *et al*. (2015)[20]: "Embryo culture conditions *in vitro* are likely to be suboptimal compared to those *in vivo*."

Is this just rhetoric or a reasonable consideration?

Describing *in vitro* data as the "best available" is a weak claim in the absence of equivalent natural *in vivo* data. The extent to which *in vitro* embryos are representative of *in vivo*

embryos is precisely the point in question. Is there really numerical consistency between natural and IVF/ART embryos? There may be consistency in sex ratios[21], but does that extend to aneuploidy rates, mosaicism, epigenetic defects, implantation potential, spontaneous abortion rates, etc? These are big questions and this article is not the place to answer them. However, if 70% loss[14] is the natural benchmark by which IVF/ART embryos are judged to be equivalent to natural embryos[22], but the true rate of natural loss lies in the range 40-60%, this therefore casts doubt on the judgement that IVF/ART and natural embryos are equivalent. Furthermore, the suggestion that IVF/ART and natural embryos may be different is neither radical, novel, nor strong[23]. However, the real reason I do not consider IVF/ART embryo data is that the article is a critique of data from natural circumstances. Comparison of natural and IVF/ART embryos is a project for the future.

The reviewers refer to my "tenuous or irrelevant reasons" why ART embryos are not useful for quantifying early embryo mortality, yet they provide the perfect reason themselves: "it is true that ART embryos are different from natural embryos in ways that could influence an estimate of fetal wastage"[24]. Nevertheless, I do discuss circumstances in which different ART interventions (e.g., observation of *in vitro* fertilisation *per se*; retrieval of embryos following timed artificial insemination, as well as AID/IVF success rates) may cast light on embryonic/fetal wastage.

Orzack & Zuckerman extrapolate from 84,881 aneuploidies among 139,704 IVF/ART embryos[21] to an "implied percentage of fetal wastage" of 60.8%. They state that this is the "central tendency" of "many reports" that I dismiss as unreliable. Of course, if this were true, then the observation would add little to what was already known. It is not clear which are the "many reports".

Let us consider the hypothesis that *in vitro* aneuploidy predicts natural total fetal wastage. Firstly, "The only well-established epidemiological facts about EPL {*early pregnancy loss*} are that about 50-60% of cases are associated with a chromosomal defect of the conceptus"[25] suggesting that euploid embryos may also fail. Secondly, "FISH may overestimate the incidence of aneuploidy"[21,26] suggesting a proportion of apparently aneuploid embryos may not fail. Furthermore, aneuploidy may not developmentally compromise embryos[27]; estimates of IVF/ART embryo aneuploidy/mosaicism vary considerably[28]; mosaic embryos can self-correct[29]; aneuploidy in trophoblast/placental cells may be less developmentally problematic[23] – who knows, it may even be advantageous!

The point is simple. There are too many undefined variables associated with IVF/ART embryos to shed more than the faintest light on the question of natural embryo survival. I have included a brief discussion of some of these issues and edited the penultimate paragraph to be more circumspect by replacing an "are" with a "may be". I hope this meets with the reviewers' approval.

Gray Literature

On several occasions, the reviewers refer to Gray Literature. They offer a revealing account and speculate on its continuing 'life'.

Gray Literature has been defined as follows: "That which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers."[30,31]

The list of references reproduced by the reviewers, starting with Opitz, 2002 and ending with McCoy *et al*. 2015 are all from academic books, journals, or text books. They are all published by commercial publishers. They were all written (with one exception) by medical practitioners or scientists, many of whom are experts in reproductive biology. The one exception (Harris, 2003) is a moral philosopher; however, the reviewers usefully point out that his estimate comes from a well-known and eminent reproductive biologist.

None of this is Gray Literature. *Human Reproduction Update*, *Fertility & Sterility* & *PLOS Genetics* are reputable academic journals. Many of these articles will have been peer-reviewed. Even pieces "akin to newspaper articles" (the Drife (1983) *BMJ* piece could be described as such and was probably not peer-reviewed[32]) are subject to editorial control, and an expectation of academic professionalism is surely reasonable from such experts.

The reviewers state that it would "be preferable if attributions were better and speculation was better highlighted". I agree. Yet they highlight my 'so-called' "errors of analysis" and "overstatement" whilst passing over errors and overstatement in these citations as "nothing inherently problematic".

What Orzack & Zuckerman describe and defend is not Gray Literature, but 'Gray Scholarship'.

Heuristics

A heuristic estimate may be based on simplified quantitative criteria, educated guesswork, rules of thumb, common sense, past experience, etc. Despite their utility, in the absence of evidence heuristic estimates may become biased. Faced with inconsistent estimates, on the one hand, those that are heuristic or based on circumstantial evidence, and on the other, those based on well-defined analysis of relevant data, surely an appropriate scientific response is to favour the latter and re-evaluate the former.

A further problem with heuristic estimates is that the process for deriving them is not always transparent. For example, it is not obvious how Orzack & Zuckerman use the "web of circumstantial evidence" to which they refer to conclude that "human fetal wastage is likely between 50 and 75%". There is something 'gray' about this. My estimates of 10-40% preimplantation loss and 40-60% total loss are partly evidence-based and partly heuristic. They may be imperfect, and no doubt will not be the last word on the matter, but it is at least clear how they were derived[1].

Conclusion

Orzack & Zuckerman often repeat the point that my estimates are consistent with previously published values. In some cases they are, and I have drawn more attention to the fine chapter by Kline *et al*. (1989)[15] who conclude that "perhaps half of all conceptions are lost before birth"[33]. However in other cases, reported values are clearly not consistent with my estimates. I have used 70% total embryo loss as a threshold, at and above which I describe estimates as exaggerated. This is based on my previous analysis[1] and thus my claim rests heavily, although not solely, on its credibility. There are other reasons to cast doubt on these high values, but these are for another time. I have modified the conclusion of my article to highlight that while precision may be elusive, exaggeration can be avoided.

"Nature abhors a vacuum", so the proverb says, but how science, or more properly scientists, should fill it is another matter entirely. Recognising and quantifying limits of knowledge is an essential part of a credible scientific process. As a philosopher once wrote: "Wovon man nicht sprechen kann, darüber muss man schweigen"[34].

References:

1. Jarvis GE: **Estimating limits for natural human embryo mortality [version 2; referees: 2 approved].** *F1000Res.* 2016, **5**:2083 (doi: 10.12688/f1000research.9479.2).
2. Anscombe GEM: **Aristotle and the Sea Battle.** *Mind*, 1956;**65**(257):1-15.
3. Oxford English Dictionary (http://www.oed.com/).
4. Roberts CJ: **Obituary: C R Lowe.** *Brit Med J*. 1994; **308**(6921): 129.
5. "It is still difficult to better the original calculations of Roberts & Lowe (1975)" from: Chard T: **Frequency of implantation and early pregnancy loss in natural cycles**. *Baillieres Clin Obstet Gynaecol*. 1991; **5**(1): 179-89.
6. Leridon H: **Intrauterine Mortality**. *Human Fertility: The Basic Components*. Chicago: The University of Chicago Press; 1977; 48-81.
7. French FE, Bierman JM: **Probabilities of fetal mortality**. *Public Health Rep*. 1962; **77** (10): 835-47.
8. Hertig AT: **The Overall Problem in Man.** In: Benirschke K, editor. *Comparative Aspects of Reproductive Failure*. An International Conference at Dartmouth Medical School. Berlin: Springer Verlag; 1967; 11-41.
9. Edmonds DK, Lindsay KS, Miller JF, et al.: **Early embryonic mortality in women.** *Fertil Steril*. 1982; **38**(4): 447-53.
10. Wilcox AJ, Weinberg CR, O'Connor JF, et al.: **Incidence of early loss of pregnancy.** *N Engl J Med*. 1988; **319**(4): 189-94.
11. Hyde KJ, Schust DJ: **Genetic considerations in recurrent pregnancy loss**. *Cold Spring Harb Perspect Med*. 2015; **5**: a023119.
12. Boklage CE: **Survival probability of human conceptions from fertilization to term**. *Int J Fertil*. 1990; **35**(2): 75, 79-80, 81-94.
13. According to Wikipedia, scare quotes too often serve to confuse rather than clarify (https://en.wikipedia.org/wiki/Scare_quotes#Criticism).
14. Macklon NS, Geraedts JP, Fauser BC: **Conception to ongoing pregnancy: the 'black box' of early pregnancy loss**. *Hum Reprod Update*. 2002; **8**(4): 333-43.
15. Kline J, Stein Z, Susser M: **Conception and Reproductive Loss: Probabilities.** *Conception to Birth. Epidemiology of Prenatal Development*. New York: OUP; 1989; 43-68.

16. Saunders P, Gibson DA. **Referee Report For: Early embryo mortality in natural human reproduction: What the data say [version 1; referees: 1 approved, 2 approved with reservations].** *F1000Research* 2016, **5**:2765 (doi: 10.5256/f1000research.9616.r19546)
17. Trounson AO. **Referee Report For: Estimating limits for natural human embryo mortality [version 2; referees: 2 approved]**. *F1000Research* 2016, **5**:2083 (doi: 10.5256/f1000research.10209.r16765).
18. Benagiano G, Farris M, Grudzinskas G: **Fate of fertilized human oocytes**. *Reprod Biomed Online*. 2010; **21**(6): 732-41.
19. Bolton VN, Braude PR: **Development of the human preimplantation embryo in vitro**. *Curr Top Dev Biol*. 1987; **23**: 93-114.
20. Bolton VN, Leary C, Harbottle S, et al.: **How should we choose the 'best' embryo? A commentary on behalf of the British Fertility Society and the Association of Clinical Embryologists**. *Hum Fertil (Camb)*. 2015; **18**(3): 156-64.
21. Orzack SH, Stubblefield JW, Akmaev VR, et al.: **The human sex ratio from conception to birth**. *Proc Natl Acad Sci USA*. 2015; **112**(16): E2102-11.
22. Niakan KK, Han J, Pedersen RA, et al.: **Human pre-implantation embryo development**. *Development*. 2012; **139**(5): 829-41.
23. Ledbetter DH: **Chaos in the embryo**. *Nat Med*. 2008; 14(5): 490-1.
24. Zuckerman JE, Orzack SH. **Referee Report For: Early embryo mortality in natural human reproduction: What the data say [version 1; referees: 1 approved, 2 approved with reservations]**. *F1000Research*. 2016, **5**:2765 (doi: 10.5256/f1000research.9616.r22945).
25. Jauniaux E, Burton GJ. **Pathophysiology of histological changes in early pregnancy loss**. *Placenta*, 2005; **26**: 114-23.
26. Treff NR, Levy B, Su J, et al.: **SNP micro-array based 24 chromosome aneuploidy screening is significantly more consistent than FISH**. *Mol Hum Rep*. 2010; **16**(8): 583-9.
27. Bolton H, Graham SJL, Van der Aa N, et al.: **Mouse model of chromosome mosaicism reveals lineage-specific depletion of aneuploid cells and normal developmental potential**. *Nat Commun*. 2016; **7**: 11165.
28. Vanneste E, Voet T, Le Caignec C, et al.: **Chromosome instability is common in human cleavage-stage embryos**. *Nat Med*. 2009; **15**(5): 577-83.
29. Munne S, Velilla E, Colls P, et al.: **Self-correction of chromosomally abnormal embryos in culture and implications for stem cell production**. *Fertil Steril*. 2005; **84** (5): 1328-34.
30. The Grey Literature Report (http://www.greylit.org/about).
31. See also: Wikipedia (https://en.wikipedia.org/wiki/Grey_literature); GreyNet International 1992-2017 (http://www.greynet.org/home/aboutgreynet.html); California State University, Long Beach Libraries (http://csulb.libguides.com/graylit).
32. Personal Communication from *Brit Med J*, 18th April 2016.
33. Kline J, Stein Z, Susser M: **Developmental Abnormalities: I. Measuring Frequencies.** *Conception to Birth. Epidemiology of Prenatal Development*. New York: OUP; 1989; 69-80.
34. Wittgenstein L: **Tractatus 7** *Tractatus Logico-Philosophicus*. New York: Harcourt, Brace & Co, Inc.; 1922 (translated as: *Whereof one cannot speak, thereof one must be silent*).

Competing Interests: No competing interests were disclosed.

Reviewer Report 09 February 2017

https://doi.org/10.5256/f1000research.9616.r19502

Jack Wilkinson
Centre for Biostatistics, School of Health Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre (MAHSC), University of Manchester, Manchester, UK

Thanks for the opportunity to review this high-quality manuscript. Peer review can be a chore, but this was a pleasure to read.

I will state that my training is in statistics and research methodology. Although much of my work is in the field of fertility, I have no clinical expertise and no familiarity with the literature discussed in this review. Any comments I make are from the point of view of the statistician and, with respect to the subject-matter, the layperson.

I am unable to comment on whether or not the body of evidence discussed in the review is comprehensive. However, the critical appraisal of these studies is conducted to a high standard, with a strong command of quantitative research methods on display. I can't fault it. The reader is left in no doubt as to the considerable limitations (many of which appear to be fatal) of these studies. All data used in the manuscript have been made available for the purposes of reproducing the analysis.

I was slightly confused by the description of the simulation study as a two-stage procedure in the critique of Roberts & Lowe. If I understand correctly, sets of simulated values for five quantities were drawn from Normal distributions centred around the estimates used by Roberts & Lowe, with standard deviations equal to these values multiplied by 0.2. Each time a new set of these five quantities was drawn, the values were used to calculate (predict) a value for embryo loss. This was done 100,000 times. However, the author speaks of 1,000 simulations, each containing 10,000 separate estimates. It is unclear what exactly varied within and between the 1000 simulations. If the data generating model was the same for all of these (ie: this was just done for computational reasons), then it would be helpful if the author could make this clear in the text.

The author assumed that the simulated quantities were independent in the simulation – I confess to having no real intuition as to the implications of this assumption. However, I don't believe this would affect the author's conclusion.

One minor typo; 'this is far from being a robust pregnancy diagnosis and in different study [46]...'

I believe that it would be appropriate to accept this manuscript without revision, although the author may wish to clarify the point about the first simulation described above.

Competing Interests: Conducting peer review may be beneficial to my career. I am funded by a Doctoral Research Fellowship from the National Institute for Health Research (DRF-2014-07-050). The views expressed in this peer review are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health. JW is a statistical editor of the Cochrane Gynaecology and Fertility Group.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 30 Apr 2017

Gavin Jarvis, University of Sunderland, UK

I would like to thank Dr Wilkinson for his helpful review. I hope the following clarify the points raised.

Roberts & Lowe Simulation: 10,000 simulated records provided data to generate 1 each of the following parameters: mean, median, and 2.5^{th} & 97.5^{th} percentiles. Repeating this process generated slightly different values for these parameters owing to the specification of the random number generator. Hence, the "10,000 simulations step" was repeated to obtain 1,000 means, medians and percentile values. It is the means of these parameters that are reported. The 100,000 simulated records were generated separately, albeit using the same model structure, simply to produce data for the Figure. I have edited the text to clarify these points.

Independence of random variables: It is plausible to suppose that some of the random variables may be correlated, e.g., length of cycle and length of fertile period. However, they may not be, or the extent of any correlation may be weak. I do not believe that constructing a full variance-covariance matrix for the model would shed any further light on the precision of the estimates of embryo mortality. There are too many undefined and imprecise variables for the model to be quantitatively useful. A solution to this imprecision would be to obtain and use more robust estimates, rather than build a more complex statistical model.

The minor typo has been corrected. Thank you.

Competing Interests: None

Reviewer Report 20 January 2017

https://doi.org/10.5256/f1000research.9616.r19546

?

Philippa Saunders
MRC Centre for Inflammation Reserach, University of Edinburgh, Edinburgh, UK

Douglas A Gibson
MRC Centre for Inflammation Research, University of Edinburgh, Edinburgh, UK

The author has provided have provided a provocative and timely review of the evidence related to pregnancy success. The author has focused on the evidence as he sees it from four different categories of review published in the last few decades.

We think the title is appropriate and will attract attention and the abstract is generally well drafted but the final sentence ends rather abruptly. We suggest the author might wish to consider a more robust/informative ending to his abstract as this will be read alone in Pubmed.

The following comments are provided in a spirit of trying to increase access to this article for a broader readership than might not be otherwise able to consider its contents, i.e. as currently written it seems largely to appeal to a people who might be interested in statistical analysis. Specifically we would like to see the author consider see how he might frame the evidence he provides alongside a timeline of the different stages of early pregnancy – this would mean individuals who are not well versed in reproductive function would be able to understand the arguments he is providing.

We are pleased to see this article being written. We think it is timely, thought-provoking and this is an excellent moment in which to consider in realistic terms the kind of evidence that is constantly requoted in the debate about how fertile the human species is. Currently this topic is dominated by data from studies on women who are sub/infertile receiving medical support to achieve a pregnancy.

Specific points

1. Who is the audience for this paper and does the introduction set the scene in such a way that the reader will be both interested and motivated to read the remaining part of the paper, which I would like to see them do? I think as written the Introduction may not achieve this objective. For example the first sentence starts with some glib comments about it being 'widely accepted' that under natural circumstances human embryo mortality is high, and then there is an extensive section quoting number of popularist articles and websites – why have this up front? It seemed to undermine the erudite arguments of the rest of the paper.

2. The second paragraph with some modification would make a sufficient introduction. The aim of the review as stated in the discussion 'How many fertilized human embryos..?' should also be frontloaded at some point here. Clearly embryo mortality is of interest to both

reproductive biologists and fertility doctors but why not also mention couples trying to conceive?

3. Reading the introduction we were struck by the pressing need for 'key terms' box – the kind of thing you see in Nature papers – where there is a definition of each of the terms used, e.g. Fecundability, embryo, HCG, etc. If this paper is going to be read by individuals who are not fertility experts or experts in reproductive biology but people interested in ethics or chance or statistics, I think they will be very confused by the different terms that are used.

4. What is not clear from the paper is the chronology of the observations/data being discussed. It is common for people (even those familiar with the field but who work on animal models) to be very confused by the timings in women. For example – the day on which fertilisation takes place versus the last menstrual period, e.g. fertilisation versus gestation versus the first day (depending on when you count from) on which you might reasonably expect to detect HCG in the urine. We would argue there needs to be a figure defining when each of these happens in terms of days in a woman's reproductive span. This could also help clarify the points in the process that the probabilities of $\pi_{FERT,}$ π_{CLIN} etc can apply to.

5. The second piece of information where we think it would be very helpful is under the section called 'What the data say' where the terms such as 'old' are added and there are no dates or refs provided. What do they mean by 'old' - pre 1960, pre 1950, pre 1940?

6. Because the author has used numbered references, there is also no sense of the relationship of one study to another in terms of dates i.e. how they chronologically relate to each other. Some minor reworking in which the author says, for instance, "the work of Hertig and Rogg in the 1950's" would be helpful.

7. The author is also slightly confusing when talking about the pregnancy study (ref 42) not giving the names of the authors nor the date on which it was published in the section on page 4, and then in the reference, for instance in Fig 2, they talk about the pregnancy study ref 42 but in the figure it is shown as French and Bierman 1962. This is the kind of things that make it difficult to get a sense of the chronology of observations and how people have built on each other's observations in order to support subsequent studies, and this after all is one of the most crucial points of this paper.

8. On page 6 we finally get to some discussion about modern pregnancy tests. It is not until some pages after that we know whether they are in blood or urine. Mid cycle elevation of HCG - this is not defined in terms of days (cf comments above). For information the fact that these assays were likely to be urine-based assays is not mentioned until page 7.

9. We think many aspects of this paper are extremely well argued, very much so the data provided. The very great detailed analysis in Table 3 and also in other parts of page 7, and some very good points are made about the over-emphasis on using data from patient groups where infertility is probably one of the reasons for presentation that may have caused a less robust data set.

10. The author makes a valid argument about potential subfertility within the Hertig cohort but

this is not balanced. Equally, these women were selected for proven fecunditity and this factor affects interpretation of this cohort as much as the other.

11. On page 10 the discussion starts with a key question how many fertilised human embryos die. It is slightly frustrating that this was not put up front as the question being addressed in this paper. Maybe the author might like to consider setting out aims more clearly.

12. Again, in the discussion, many of the arguments being made would have been greatly enhanced by telling us the dates on which some of these studies were conducted. When looking at the reference list I see many of them were in the '80s and early '90s.

13. We wonder if the first paragraph on page 12 might reasonably be eliminated - it feels repetitive compared to other parts of the paper. I think the discussion of the studies by Macklon review ref 20 is extremely insightful and useful. However we draw the author's attention to a more recent study by Macklon and Brosens which we believe puts forward some interesting arguments that might reasonably be discussed in his study about how the endometrium in which the embryos are set to implant might be acting as a 'sensor' of embryo quality. This is in Biology Reproduction 2014, vol 91. There is also a complementary paper in Sci Rep, vol 6, Brosens *et al.* 2014.

14. The conclusion of the discussion seems more like a continuation of the critique of the final few paragraphs. It would be desirable to provide a concluding paragraph which holistically draws together the content of the review. Again the heavy use of quoting references as appears in the introduction masks the opportunity for the author to provide his own conclusions.

In summary we welcome this review which we think makes many erudite comments on a difficult field.

Competing Interests: No competing interests were disclosed.

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Author Response 30 Apr 2017

Gavin Jarvis, University of Sunderland, UK

I would like to thank Professor Saunders and Dr Gibson for their helpful review. I have tried to address each point as detailed below.

The abstract has been re-drafted to provide a clearer and more robust conclusion. I have also brought the "How many..." question posed in the discussion to the beginning of the Abstract and into the Introduction as suggested.

I hope that the changes made will make the article more accessible to a wider readership.

Responses to specific points

1. The article is intended primarily for a scientific audience, although I agree that it may have wider appeal and I would be pleased for it to be read as such. The references to claims in the press and popular media are important (see later); however, I agree that as an opening this may be a distraction. I have therefore moved this section to the discussion. The list of scientific citations that make claims about embryo mortality is essential since it substantiates my claim that high embryonic mortality is widely reported. Additionally, it illustrates the large variance in these estimates.
2. The introduction has been re-organised along the lines suggested. The second paragraph has become the start and the quantitative claims incorporated into it. The importance of embryo mortality to women trying to conceive is clear, and was already present in Version 1. It remains in Version 2 and has also been emphasised in the discussion. The “How many…” question has been frontloaded as suggested, and is now also in the abstract.
3. A glossary of key terms box has been provided.
4. A new figure provides a timeline for non-fecund and fecund cycles and key biological events.
5. Most of the pre-1960 references are found in databases, e.g., PubMed; however, search terms appear less comprehensive as articles get older. Hertig (1959) (Ref. 90, PMID: 13613882) has no reference to pregnancy loss in its MeSH terms. Neither Rock & Hertig (1942) (Ref. 92) nor Opitz (2002) (Ref. 29) is indexed in PubMed. A PubMed search on {"early pregnancy loss"[All Fields]} generates 831 hits (25th April 2017) of which the earliest is 1971. However, I agree that the description in Version 1 is vague. I have removed the reference to “old”.
6. I have edited the text in several places to include names and dates. I hope this will assist readers in following the chronology. The hCG studies are already arranged by date in Table 3 and Figure 3.
7. Reference numbers have been added to the legend within the Figure, and names and dates included in the legend text to increase clarity.
8. All of the 13 studies measured hCG in urine except two. These two employed serum samples and have been identified in the text. The term “mid-cycle” has been removed and the meaning clarified using the new Figure 1.
9. Thank you for these comments.
10. I agree with the point made. In this section the first point I make about the 210 recruited women is that they were of “proven fertility”. The same point is made again in the commentary on Hertig’s data. I have also edited the text in response to Reviewer 3 and hope that the final result is appropriately balanced.
11. The question (slightly modified) has been included in the Abstract and Introduction. I hope that this helps to clarify and reinforce the purpose of the article.
12. Some dates have been incorporated to enhance chronological clarity (see point 6).
13. The first paragraph on page 12 addresses the importance of biological variance. It does not go into detail but stresses that point estimates of risk do not provide the whole picture when considering either populations or individual cases. As I have put it, a neglect of variance fosters a misleading appreciation of reality. I would prefer to retain this paragraph, in the hope that it will encourage readers to consider the importance and implications of numerical diversity when interpreting data.The

arguments proposed by Macklon and Brosens relating to endometrial receptivity are indeed interesting. However, they propose mechanistic explanations for implantation failure and do not directly address the issue of how frequently such events occur. Nevertheless, their inclusion is contextually valuable, and I have made some comments on their studies.

14. The final paragraph has been edited. The quotations are useful to make it clear that I am not alone in drawing attention to the limitations of the available data. I have endeavoured to summarise the broad purpose and value of this work.

Competing Interests: None