Research Methodology in Recurrent Pregnancy Loss

Ole B. Christiansen, MD, DMSc[a,b,*]

KEYWORDS

• Recurrent pregnancy loss • Recurrent miscarriage • Medical research
• Research studies • Methodological flaws

KEY POINTS

- There is currently substantial disagreement concerning the diagnostic criteria for recurrent pregnancy loss (RPL), which renders comparisons between research studies in the area difficult.
- There are numerous methodological pitfalls that threaten the validity of research studies in the field of RPL and it is necessary for scientists and clinicians to be aware of them.
- Some of the methodological pitfalls are common for medical research in general, whereas some are specific for RPL research.
- Frequently seen methodological flaws in case-control studies are comparisons of biomarkers between patients with RPL and controls that differ from patients with regard to previous ongoing pregnancies, relevant endocrine factors, and viability of the fetal tissue at the time of sampling.
- In cohort studies, incomplete follow-up of patients in many studies has resulted in huge variations in estimates of the prognosis after RPL.
- Only a few small and heterogeneous double-blinded placebo-controlled trials of treatments of RPL have been carried out with very heterogeneous results.
- Proposals are given for improvements in the design of research studies in RPL that hopefully can improve the quality of studies in the future.

INTRODUCTION

Compared with the situation in other reproductive medicine disorders, such as tubal or male factor infertility and in other areas of medicine, there is very little consensus about which investigations are useful for identifying causes or estimating the prognosis and which treatments are effective in recurrent pregnancy loss (RPL). It is generally agreed that when the tubes are occluded, as diagnosed by laparoscopy or

[a] Fertility Clinic 4071, Copenhagen University Hospital Rigshospitalet, Blegdamsvej 9, DK-2100, Copenhagen, Denmark; [b] Department of Obstetrics and Gynecology, Aalborg University Hospital, Reberbansgade, DK-9000, Aalborg, Denmark
* Fertility Clinic 4071, Copenhagen University Hospital Rigshospitalet, Blegdamsvej 9, DK-2100, Copenhagen, Denmark.
E-mail address: olbc@rn.dk

Obstet Gynecol Clin N Am 41 (2014) 19–39
http://dx.doi.org/10.1016/j.ogc.2013.10.001 **obgyn.theclinics.com**
0889-8545/14/$ – see front matter

hysterosalpingography (HSG), pregnancy can happen only after in vitro fertilization (IVF), and when the number of viable spermatozoa is very low, pregnancy can happen only after intracytoplasmatic sperm injection (ICSI). It is also generally agreed that IVF or ICSI are very efficient treatment methods for the 2 reproductive disorders. In recurrent early pregnancy loss there is much more disagreement about diagnosis, cause, and treatments. Although most guidelines from specialist societies do not support the screening of RPL women for hereditary thrombophilia factors or peripheral blood or endometrial natural killer (NK) cell numbers and function,[1–3] many clinics are still doing this, and whereas the Cochrane review[4] or national guidelines do not recommend immunotherapy[1,2] or preimplantation genetic screening (PGS)[5] for RPL, immunotherapy and PGS are still widely used in many clinics.

There could be many reasons why doctors very often do not adhere to the clinical guidelines regarding RPL:

- Pressure from desperate patients to do something although very few proven therapies really exist
- The doctors' economical motives, as many patients are desperate and willing to pay a lot of money for treatments that may provide them with some hope for a solution to their problem
- Current guidelines are based on few, small, and often poor-quality studies that cannot support strong, evidence-based recommendations

In this author's view, the third statement is the most important cause for this poor adherence to RPL guidelines. With regard to almost every diagnostic test or treatment for RPL, it is possible to find studies presenting data strongly in favor of this test or treatment and other studies strongly against. It is therefore often up to clinicians themselves to decide which studies they find trustworthy.

The aim of this article is to highlight pitfalls in research methodology that may explain why studies in RPL often provide very divergent results, and it is hoped that insight in this issue may help clinicians to decide which published studies are most valid. It may help researchers to eliminate methodological flaws in future studies, which may hopefully come to some kind of agreement about the usefulness of diagnostic tests and treatments in RPL.

CONTROVERSIES OF DEFINITION

It is disputed how to define RPL. It is important to realize that RPL is defined quite differently from most other diseases. Most diseases are defined by some unique pathoanatomical, clinical, or paraclinical findings being permanently present, whereas RPL is defined by a series of transient events in the past that may have be poorly registered.

The controversies concerning how to define RPL deal with

- The number of miscarriages needed for the diagnosis
- The role of nonconsecutive miscarriages
- The role of preclinical losses

Until 10 years ago, the definition of RPL was undisputedly 3 or more consecutive miscarriages, because it was commonly agreed that after 3 miscarriages the chance of live birth the next pregnancy without treatment was substantially decreased.[6] However, during the recent years, some national guidelines have adopted an RPL definition of only 2 clinically recognized miscarriages[1] or 2 not necessarily consecutive miscarriages.[7] This redefinition is based on finding similar frequencies of selected factors suggested to cause RPL: uterine abnormalities; antiphosphospholipid antibodies

(APL); parental chromosome aberrations; or the factor V Leiden mutation in women with 2, women with 3, and women with more miscarriages.[7,8] It is argued that when such risk factors already recommended in the screening of couples with 3 miscarriages can be found with similar prevalence in those with 2 miscarriages, they should also be examined in the latter. If the same tests are recommended in couples with 2 as well as those with 3 or more miscarriages, it is a short step to redefine RPL as 2 or more miscarriages. There is also now disagreement regarding which kind of pregnancy losses should be included in the criteria for RPL. Thirty years ago, because of the nonexistence of ultrasonic examinations and high sensitive pregnancy tests, pregnancies could not be diagnosed before gestational week (GW) 6 to 7; therefore, the pregnancy losses considered in the RPL diagnosis were miscarriages, which had normally been confirmed by curettage and histology. Pregnancies can now, owing to highly sensitive and specific β-human chorionic gonadotropin (hCG) tests, be diagnosed a few days after the due menstrual period, and many of these (biochemical pregnancies) will fail before it is possible to do transvaginal ultrasound. There is thus an urgent need to find a place for these kinds of losses in the RPL diagnosis. Because transiently positive pregnancy tests at the time of the due period are a frequent finding in women not using anticonception,[9] many gynecologists have been reluctant to include biochemical pregnancies in the RPL diagnosis, and the American Society for Reproductive Medicine (ASRM) definition of RPL (2 or more clinical miscarriages) completely disregards them.[1] A recent study from the European Society for Human Reproduction and Embryology (ESHRE) early pregnancy special interest group on the other side found that in patients with RPL, each early pregnancy loss confirmed only by a β-hCG test displays a negative prognostic impact equal to that of a clinical miscarriage, supporting the view that biochemical pregnancies should be included in the RPL diagnosis.[10] The different diagnostic criteria recommended in the national guidelines or by leading RPL clinics are[1,2,7,10] as follows:

- ≥3 consecutive pregnancy losses before 24 weeks of gestation
- ≥2 consecutive clinical pregnancy losses
- ≥3 consecutive clinical miscarriages and biochemical pregnancies
- ≥2 not necessarily consecutive pregnancy losses before 24 weeks of gestation

In numerous studies, it was found that the strongest predictive factor for new miscarriage in patients with RPL is the number of previous losses.[6,11] Because the same diagnosis, RPL, for the time being covers patients with a wide range of previous losses and therefore very different pregnancy prognoses, it will in the future be increasingly difficult to compare and combine different studies of outcome in patients with RPL (eg, results of randomized controlled trials [RCTs]). The only way to overcome this obstacle is that the investigators in such studies stratify the results according to the number of previous clinical miscarriages, as well as biochemical pregnancies or pregnancies of unknown location (PULs).

TYPES OF RESEARCH STUDIES

Different pitfalls characterize the 3 main types of research studies done in RPL:

- Case-control studies
- Cohort studies
- Intervention and treatment studies

For each category of studies, I provide an overview of the pitfalls that threaten the validity of the studies and the flaws often seen in publications: some of them can be

seen in other areas of medical research and they are discussed superficially, whereas others that are specific for RPL research are discussed in more detail.

CASE-CONTROL STUDIES

Case-control studies always have a retrospective design. The frequency of a potential risk factor is investigated in a group of patients who have been sampled during a defined period (typically in a single clinic) with a specific disease diagnosis and compared with the corresponding frequency in a group of randomly selected individuals either without the disease or (if the disease is rare) selected from the background population. An estimate of the potential risk factors' association with the disease is typically given by the odds ratio (OR) with 95% confidence limits indicating the ratio between the frequency of the risk factor in diseased individuals and in controls.

Methodological errors in case-control studies can occur during the sampling of both patients and control subjects and errors can occur in the testing for potential risk factors.

Flaws in Sampling of Patients and Controls

Inconsistent diagnosis

The different definitions of RPL have already been discussed. Different RPL definitions will make comparisons between case-control studies originating from regions with different definitions increasingly difficult in the future.

Misclassification of disease/outcome status

Even if identical RPL definitions are used, patients in different studies can differ regarding the severity of the disease: the number of previous miscarriages. Because the RPL diagnosis is based on a series of past events, the validity of the diagnosis is dependent on the quality of the information that is available about these events. In many clinics, information about previous pregnancy losses comes primarily from interviewing the patients. Misclassification of disease/outcome status (the number of previous miscarriages) can be random or nonrandom. There are many examples of random outcome misclassification when dealing with miscarriages. Only 71% of miscarriages reported by women without RPL who have been treated at hospital could be verified in hospital records,[12] and in a retrospective study, 348 women recalled 30 (6%) of 507 miscarriages that were not reported in a prospective study several years before.[13] Random outcome misclassification results in an underestimate of the hypothesized association between exposure and disease/outcome in case-control studies (**Table 1**).

Retrospective information about previous pregnancy losses can also be subjected to nonrandom misclassification, also called recall or information bias. This misclassification is a difference in the ability or inclination to remember or report events or exposures in the past in individuals with or without particular characteristics. Women who had given birth to a child with a congenital malformation will search their memory extensively for any potentially teratogenic exposures during pregnancy and therefore retrospectively report more exposures than women who had delivered a healthy child, although there is no real difference in the frequency of harmful exposures in the 2 groups. Some women with 1 or 2 confirmed miscarriages will be more prone to interpret delayed menstruations or recall previous terminations as miscarriages than women without recent pregnancy losses, and thus erroneously get a diagnosis of RPL. Such nonrandom misclassification of disease/outcome status will result in either an overestimation or underestimation (see **Table 1**) of the true size of the hypothesized

Table 1
Important methodological pitfalls typical for research studies in RPL

Factor to Evaluate	Effect on Study Outcome
Definition of RPL ≤2 miscarriages	Decreases difference between risk variables in CCS and treatment effect in RCTs
Random misclassification	Decreases difference between risk variables in CCS
Nonrandom misclassification	Decreases or increases difference between risk variables in CCS or outcome variables in CS
Ascertainment bias	Increases difference between risk variables in CCS
Relevant mismatches between patients and controls	Increases difference between risk variables in CCS
Lack of protocol details/multiple testing	Overestimates significance of chance findings
Historical controls	Increases effects in treatment studies
Nonblinding	Increases or decreases treatment effects in RCTs
Premature termination after interim analysis	Decreases treatment effect in RCTs
Inclusion after detection of fetal heart action	Decreases treatment effect in RCTs
Poor characterization of RPL and subgroups of RPL	Renders comparisons between CCS and RCTs difficult and makes meta-analysis difficult
Unfounded exclusions of RCTs in systematic reviews	Bias combined risk estimates in meta-analyses

Abbreviations: CCS, case-control study; RCT, randomized controlled trial; RPL, recurrent pregnancy loss.

relationship between exposure and disease/outcome (RPL). To avoid random and nonrandom misclassification in RPL case-control studies, information about previous pregnancies should as much as possible be confirmed from external sources: records from hospitals, fertility clinics, and practitioners and serum hCG measurements should be documented from laboratory reports.

Ascertainment bias

In case-control studies, the frequency of a potential risk factor for the disorder under study (eg, RPL) is compared between patients and controls. Ascertainment or selection bias happens when patients with some clinical or paraclinical risk factor are preferentially referred to a specific clinic because of knowledge of the clinic's expertise or interest, and a study focusing on this particular risk factor is undertaken in the clinic. An example of ascertainment bias in RPL has been reported by Out and colleagues.[14] In this study, the frequency of the APLs anticardiolipin and lupus anticoagulant was higher in patients with RPL referred to a Dutch center for APL research than in controls. However, when patients with RPL with a history of thromboembolic or lupuslike symptoms were excluded, the prevalence of APL in the remaining patients with RPL did not differ from that of controls. The high prevalence of APL in the total group of patients resulted from the preferential referral of patients with RPL with APL-associated symptoms to the clinic due to its special expertise (see **Table 1**).

Ascertainment bias can also work in controls. In studies of risk factors for adverse pregnancy outcome, information is often obtained or blood samples drawn from women with healthy pregnancies coming for routine pregnancy control. Women giving

informed consent to become controls are often better educated than average women, thereby decreasing the occurrence of lifestyle factors and exposures that may be harmful in pregnancy. When staff members or healthy blood donors are used as controls, these also undergo positive selection for being healthier than the average population.

Flaws in Estimating Risk Factors

Misclassification of exposure status

If the risk factor(s) under study is a potentially harmful exposure (eg, infection, smoking, medicamentation) in which the estimate of exposure is based on information retrospectively obtained from the patients and controls themselves; as previously discussed, this information can be subject to both random and nonrandom misclassification. If exposures have the same probability of being overreported/underreported in patients and controls, we are dealing with random misclassification and this will lead to an underestimate of the hypothesized relationship between exposure and disease/outcome. When the probability of exposure misclassification differs between patients and controls, nonrandom misclassification occurs. An example of nonrandom misclassification is when patients with RPL often recall an episode of fever during a pregnancy that subsequently failed, whereas controls with successful pregnancies report such episodes less often, although the incidence of fever in the 2 groups may be similar. Nonrandom misclassification of exposures can lead to an underestimate or overestimate of the true size of association between exposure and disease/outcome. To avoid misclassification in RPL case-control studies, data about exposures should be collected as much as possible from external sources: records from hospitals or practitioners or be collected before the outcome is known: in early pregnancy before miscarriage is diagnosed.

Confounding

In case-control studies, a confounding factor is a clinical or paraclinical factor that is associated with both the risk factor and the disease/outcome under study. If adequate measures are not taken, a confounding factor can be mistaken as a causal factor or diminish the estimate of the impact (OR) of the risk factor under study. In RPL studies, age is an important and common confounding factor. In studies of the prevalence of autoantibodies in patients with RPL and controls, age is associated with both the risk of miscarriage and RPL and the occurrence of autoantibodies. Elimination of the confounding effect of age during the inclusion phase of such studies can be undertaken by age-matching patients and controls and in the analysis phase by reporting autoantibody frequencies in different age strata of patients and controls and subsequently do adjustment by multivariate statistical methods.

Mismatch between patient and controls group

In case-control studies, the aim is to compare diseased and healthy individuals, which is quite straightforward regarding most diseases, but in RPL, research finding a suitable control group is a much more complex task (see **Table 1**; **Table 2**). Case-control studies in RPL research typically compare the following:

- Biomarkers in the blood or endometrium of nonpregnant women or in the blood of pregnant patients with RPL with the same biomarkers in women who do not have RPL
- Biomarkers in decidual or trophoblast tissue from women with RPL who have miscarried with the same biomarkers in tissue from non-RPL women who had an induced abortion.

Table 2
Suggested optimal control groups for the investigation of biomarkers in patients with RPL or their products of conception

Nongenetic Biomarkers						Genetic Biomarkers	
Peripheral Blood		Luteal Phase Endometrium		Decidual/Trophoblast Tissue		Peripheral Blood/ Tissue	
RPL[a]	Controls[a]	RPL[b]	Controls[b]	RPL[c]	Controls[c]	RPL	Controls
Primary/NP	Nulligravida/NP	Primary	Nulligravida	Primary, Euploid male embryo	Primary RPL Aneuploid embryo	All	Multipara
Secondary/NP	Previous birth/NP	Secondary	Previous birth	Secondary, Euploid male embryo	Secondary RPL Aneuploid embryo		
Primary GW5	Primigravida GW5						
Secondary GW5	Previous birth GW5						
Primary > GW6 Euploid male embryo	Primary RPL > GW6 Aneuploid embryo						
Secondary > GW6 Euploid male embryo	Secondary RPL > GW6 Aneuploid embryo						

In each double column, the suggested optimal controls are found to the right of the RPL column.

Abbreviations: GW, gestational week calculated from the first day of last menstrual period; NP, not pregnant; RPL, recurrent pregnancy loss.

[a,b,c] Women in each comparable pair should ideally have similar estrogen and progesterone levels in the nonpregnant state and similar hCG, estrogen, and progesterone levels in the pregnant state.

In studies of external exposures or genetic biomarkers, the ideal control group for patients with RPL is women with proven fertility: typically women with 2 or more uncomplicated births and no miscarriages, as genetic polymorphisms will not change according to reproductive history or be affected by endocrine factors or inflammation (see **Table 2**).

However, it is much more complex to identify the ideal controls for the investigation of nongenetically determined biomarkers in the blood, the endometrium, or trophoblast tissue. Some biomarkers, although exhibiting no impact on pregnancy outcome, may be different in patients with RPL and women with previous births merely because of their different reproductive histories.

An illustrative example is anti-HLA antibodies and studies of their role in RPL. Several years ago many articles reported that these antibodies could be detected much less frequently in the blood of patients with primary RPL (no previous ongoing pregnancies) than in control women who had previously given birth. It was postulated that lack of these "blocking antibodies" was a cause of the RPL and that deliberate immunization of the patients with paternal lymphocytes with the aim to stimulate antibody production would improve the pregnancy prognosis. It is now generally recognized that these antibodies are a common feature of normal ongoing pregnancy due to passage of fetal cells through the placenta into the maternal circulation during late pregnancy and they often persist for many years.[15] Therefore, it is not surprising that patients with primary RPL normally lack these antibodies, whereas multipara are often positive.

It has also been shown that cellular immunity is being permanently changed after a birth. Clones of cytotoxic lymphocytes with specificity for male-specific minor histocompatibility (HY) antigens develop in half of the women pregnant with a male fetus and the anti-HY cytotoxicity remains unchanged up to 18 years after delivery.[16] Long-term persistence of regulatory T cells[17] or/and persistence of fetal cells in the maternal circulation (fetal microchimerism) after a first ongoing pregnancy[18,19] may be the reason that most women remain immunologically tolerant to the fetus in spite of production of antibodies and lymphocytes with reactivity toward fetal antigens. The maternal immune system therefore recognizes many paternal/fetal alloantigens during an ongoing pregnancy and this very often induces permanent changes in immune reactivity to the fetus or trophoblast that may reside in lymphocytes carrying immunologic memory (memory T cells) in the peripheral blood, endometrium, or regional uterine lymph nodes. It is also possible, although much less studied, that transcription of messenger RNA (mRNA) and expression or production of proteins that are not related to the immune function (eg, receptors for hormones in the uterus, production of coagulation factors in the liver) can be permanently altered subsequent to the extensive physiologic changes taking place during a prior ongoing pregnancy (a pregnancy passing GW 22).

Because of these considerations, in RPL case-control studies controls should be matched to patients with RPL regarding a history of previous ongoing pregnancies (see **Table 2**). The ideal control group for patients with RPL who had never had an ongoing pregnancy would be women with repeated first trimester terminations due to social reasons; however, because this group is fortunately small, an alternative suitable control group would be women with no previous pregnancy. It may be argued that 1% of nulligravida would later experience RPL, but this small error is insignificant compared with the error associated with comparing nulliparous patients with RPL with multiparous controls. Patients with secondary RPL who had previously had a birth should of course be compared with controls with 1 previous live birth and no miscarriages.

Another factor that can confound case-control studies in RPL research and should be adjusted for as much as possible is differences in hormones relating to

reproduction and pregnancy. Many factors relating to immune function (eg, cytokines) and the coagulation system (eg, proteins C and S) are influenced by estrogen and progesterone levels.[20] The level of expression of a series of cytokine mRNA in endometrial cells increases markedly from the follicular to the late secretory phase, probably influenced by cyclic changes of estrogen and progesterone.[21] It has been shown that lymphocytes that can induce immunologic tolerance (regulatory T cells) are attracted to the feto-maternal interface by hCG produced by the trophoblast,[22] which may have profound importance for measurements of immune biomarkers in the uterus during pregnancy. Because concentrations of hCG, estrogens, progesterone, cortisone, and pregnancy-associated placental protein A (PAPP-A) change markedly according to menstrual cycle phase or progressing gestation, the level of biomarkers affected by hormones is dependent on phase of menstrual cycle or time of gestation. Therefore, as a main rule, nongenetic biomarkers should be investigated in patients with RPL and controls matched by menstrual cycle phase in nonpregnant women and length of gestation during pregnancy (see **Table 2**).

A further problem arises when biomarkers in the blood or uterus, which are affected by hormones, are investigated in patients with RPL just before or at the time of miscarriage and compared with similar measurements in controls with a healthy ongoing pregnancy. Most of the pregnancy-related hormones, hCG, estrogen, and PAPP-A, decrease in threatened miscarriage and finding differences in a specific biomarker between women with a miscarriage or a healthy pregnancy, respectively, may be severely confounded by the fact that hormones in the former group are lower than in the control group.

A last but important factor that may confound case-control studies in RPL research is differences in the viability and inflammatory status of the uterine content before or at the time of miscarriage in patients and at the time of induced abortion in controls. At the time of embryonic death, the trophoblast will undergo necrosis and intrauterine hemorrhage will often induce inflammation in the decidual tissue, which can be reflected in measurements of immunologic biomarkers in the blood or decidual/trophoblast tissue. In contrast, the same biomarkers in the blood of controls with an ongoing normal pregnancy or in the blood or decidual/trophoblast tissue from controls with an induced abortion on social indication will not be influenced by inflammation. Unfortunately, in many publications, levels of biomarkers associated with immune function or apoptosis (programed cell death) in the blood or decidual/trophoblast tissue are compared between women with a missed abortion and women with a normal ongoing pregnancy or women undergoing induced abortion. In many of these publications, finding differences in levels of such biomarkers are interpreted as proof that changes in maternal immune reactions or apoptosis are causing miscarriage or RPL. To avoid detecting biomarkers in patients with RPL that differ from those in controls merely due to processes being a result of rather than a cause of miscarriage, measurements in peripheral blood should be undertaken in very early pregnancy (eg, GW 5) at a time when the feto-placental unit is very tiny and not expected to affect systemic inflammatory responses and before the confounding effect of declining hormones in patients take place (see **Table 2**).[23]

Another approach to counteract the methodological problem associated with the comparison of necrotic and vital tissue is to do karyotyping of embryos from missed abortions in women with or without RPL. Biomarkers in the blood or decidual/trophoblast tissue can then be compared between patients with euploid male embryos (to avoid erroneous karyotyping of maternal tissue) and patients with embryos with a chromosome abnormality that definitively will cause early embryonic death. Levels of biomarkers that are influenced by embryonic and trophoblast necrosis and

inflammation may be equally affected in women with euploid and aneuploid miscarriages. Therefore, differences in expression of biomarkers between the women/embryos from the 2 groups can probably be attributed to factors that may have caused euploid miscarriage. Examples of such studies that point to a causal role of inflammatory cytokines in RPL is the study by Calleja-Agius and colleagues,[24] finding significantly higher plasma tumor necrosis factor-α, interferon-γ, interleukin (IL)-6, and IL-10 levels in women with euploid miscarriages than in healthy pregnancy, whereas these cytokines were not increased in women with aneuploid miscarriages. Another study similarly found increased numbers of activated leucocytes in RPL women with a miscarriage with an euploid compared an aneuploid embryo.[25]

As illustrated previously, finding a suitable control group for measurement of nongenetically determined biomarkers is a difficult task. Adherence to the recommendations given previously (see **Table 2**) may diminish the risk of conducting a case-control study that is methodologically flawed but offers no guarantee. The factors that should not differ between cases and controls in RPL research are the following:

- Number of previous ongoing pregnancies
- Levels of estrogen and progesterone in nonpregnant women
- Levels of estrogen, progesterone, and hCG in pregnant women
- Degree of viability of decidual or trophoblast tissue

I realize that very few case-control studies in RPL meet these criteria, especially the criteria of similar hormonal levels. More research should be done regarding genetically determined biomarkers because these are robust to the confounding effects discussed in this section. It is possible that some nongenetic biomarkers may also be robust to the effects of the mentioned confounding factors (eg, reproductive hormones). However, only when it has been proven in separate studies that a specific factor (eg, hormone) is not affecting the biomarker under study, a confounding effect of the factor can be ignored in subsequent case-control studies.

Lack of Protocol Data or a Priory Hypothesis

As a referee, I have often discovered that some research groups have published a series of case-control studies in different articles often in different journals, each case-control study dealing with only one specific biomarker, which was found to be significantly increased in patients with RPL. When the articles are carefully read and the data compared it seems that all the biomarkers have been investigated in the same patient and control groups in the same period. In the articles, no information was given about how many biomarkers were investigated in each study and what was the a priori hypothesis. The suspicion is that in many of these studies, the investigators have tested a huge series of biomarkers using a multiplex testing panel that can test for hundreds of (often genetic) biomarkers in 1 day and they report only data of those biomarkers that (by chance) are found significantly increased in RPL. After this "fishing expedition," the biomarkers being significantly associated with RPL are published one by one in separate articles according to the "salami method" without providing any information about how many biomarkers were investigated. In this way, the investigators can expand their publication list in an easy way. This "disguised multiple testing" is a maligned design, because the referees of the articles have no chance to know that multiple testing was done and to ask the investigators to do the appropriate statistical adjustment (see later in this article) and modify the conclusions. In the more benign cases, the investigators openly provide details of all tested biomarkers/variables in tables and the referees can then ask for adjustment if not already done.

This illustrates the problem that in many studies (especially case-control studies) dealing with RPL and other disorders, it is often unclear whether the result being reported as the main finding really was part of the a priori hypothesis from the start or whether it was a finding discovered during the conduct of the study: a post hoc finding (see **Table 1**).

As a preventive measure, high-ranking journals request the investigators to give more details of the study protocols or the protocols should be available on an online public registry (www.ClinicalTrials.Gov) before the beginning of the study. All journals should, as a minimum, request the investigators in the "materials and methods" section to provide information from the study protocol about the a priori hypothesis, main outcome measures, and sample size calculations, and to list all risk factors and biomarkers that were planned to test and actually tested.

Multiple Testing

Most case-control studies in RPL can be categorized as discovery studies in which there is no prior hypothesis about which associations are probable and therefore a series of biomarkers are investigated openly or disguised. Testing of multiple biomarkers is facilitated by the introduction of multiplex testing panels that can test hundreds of biomarkers at a time, thereby reducing the costs immensely. If, for example, 40 different biomarkers are tested in patients and controls and a *P* value for significance of less than .05 was chosen, it is expected that 2 ($40 \times 0.05 = 2$) biomarkers will be significantly increased or decreased in patients merely by chance. If no a priori hypothesis about which biomarker was the primary focus of the study was made (discovery study) the *P* values for all tested biomarkers should be subject to the Bonferroni adjustment by multiplying the *P* value with the number of comparisons made. If 1 or 2 biomarkers are still significantly associated with RPL after this adjustment, the association must be confirmed in at least 1 independent study (replication study) with the a priori hypothesis that the variables are associated with RPL.

Interpretation

Causality can rarely, if ever, be stated after finding a potential risk factor statistically significantly associated with a disease. According to Bradford-Hill criteria,[26] there are several demands for stating causality, the most important being the following:

- The criterion of temporality: that the risk factor occurs before the occurrence of the disease.
- The criterion of a biologic gradient: there should be a positive correlation between the severity of a putative risk factor and the severity of the disorder.

In studies of RPL, the first criterion can be documented only in prospective studies finding that the presence of a potential risk factor increases the risk of a new miscarriage compared with its absence. In RPL, the other criterion can be documented if there is a correlation between the severity of a potential risk factor, for example, concentration of APL antibody, and the severity of RPL = number of miscarriages in the patients. Findings in some previously quoted studies may indeed be interpreted as evidence against the investigated variables being causative for RPL because no biologic gradient was discovered.[7,8]

COHORT STUDIES

As stated in the previous section: to document that an exposure found to be associated with RPL in a case-control study is indeed a causal factor for RPL, a prospective study is needed.

A prospective study is normally designed as a cohort study. A group of individuals positive and another group of individuals negative for the exposure/potential risk factors for the disease (outcome) are identified. These 2 groups are now called cohorts and they are observed during a defined time period in which the occurrence of the outcome under study is monitored in the cohorts.

Concurrent Versus Nonconcurrernt Cohorts

In concurrent cohort studies, individuals assigned to the 2 cohorts are followed prospectively, whereas in nonconcurrent (or retrospective) cohort studies, the assignment to the cohorts is done on the basis of the detection of an exposure at a time in the past when the outcome under study had not yet happened. If the disease under study is RPL, such a concurrent cohort study could ideally assign 20-year-old women who had not yet been pregnant and who have been tested for relevant biomarkers into a hereditary thrombophilia positive and negative cohort. These 2 cohorts are then followed until the women had aged 45 years and the occurrence of RPL in the 2 cohorts is registered and compared between the cohorts. Adjustment for other risk factors for RPL, such as age at the first pregnancy attempt, body mass index (BMI), and so forth, should be undertaken in a multivariate analysis. It is clear that such a cohort study will probably never be undertaken; because of the low frequency of RPL, tens of thousands of women must be included and followed for 25 years: nobody would have the resources for that. Instead cohort studies focusing on pregnancy outcome after a diagnosis of RPL will be a easier task to perform. Still, concurrent cohort studies can take many years and therefore almost all cohort studies of patients with RPL have been nonconcurrent. To give an example, my group investigated hereditary thrombophilia factors (factor II and factor V Leiden mutations) in the patients referred to the RPL clinic in Denmark between 1986 and 2008, in 62% of the cases using DNA, which was stored but not tested for the thrombophilia-associated mutations before after the outcome of the first pregnancy after referral had been registered.[27] In most cases, the patients and their doctors were therefore unaware of the exposure status at the time of the first pregnancy after referral and none of the patients received anticoagulation treatment. In a good nonconcurrent cohort study, information about both exposure and outcome status should not be available at the time the cohorts are formed. The patients were thus in a nonconcurrent (retrospective) cohort study assigned to a hereditary thrombophilia positive and negative cohort, and outcome of the first pregnancy after referral to the clinic was “outcome.” In a logistic regression analysis, adjusting for the impact of number of previous pregnancy losses, maternal age, and smoking, it was found that the presence of the hereditary thrombophilia factors significantly decreased the OR for birth in the first subsequent pregnancy (OR = 0.48, $P = .05$) compared with their absence. Thus, a much easier approach than doing a large concurrent cohort study can provide results documenting the importance of potential causal factors in RPL.

Nonrandom Misclassification

Cohort studies are prone to methodological errors, which are important to recognize when conducting or assessing them. Nonrandom misclassification is a significant threat to the validity of cohort studies; however, unlike case-control studies in which the patients or controls are the main source of the erroneous information leading to misclassification, in cohort studies it is the researchers who are responsible. In occupational medicine, workers are exposed to substances thought to increase the risk for some disease. Because of that knowledge, these workers are often monitored more closely than nonexposed workers and symptoms of disease (outcome) may be

registered more often in the exposed than in the nonexposed cohort. Biased intensity of monitoring is also a problem in reproductive medicine research. In a cohort study of women with obesity (the exposure) or no obesity, to investigate whether obesity increases the risk of miscarriage, many obese women will have polycystic ovary syndrome (PCOS) and anovulation and therefore undergo assisted reproductive technology (ART) treatment, whereas most nonobese women will be able to conceive without the need of ART. In ART cycles, a β-hCG test is normally done 14 days after ovulation or embryo transfer, and all biochemical pregnancies will be detected, whereas among non-ART patients, they will often remain undetected. Therefore such a cohort study may show that obese women have an increased risk of pregnancy loss but this may be a result of nonrandom misclassification of the outcome.

In RPL cohort studies, misclassification is also a significant problem. The prognosis after a diagnosis of RPL is heavily disputed: various prospective studies have reported the chance of live birth in the first pregnancy after referral in patients with 3 miscarriages to be between 63% and 87%, with 4 miscarriages between 44% and 73%, and with 5 miscarriages between 25% and 52%, respectively.[11] These very different estimates are frustrating: for a patient with RPL, it is of utmost importance to know whether her chance for live birth is 44% or 73%. This variation may be caused by random misclassification within individual studies. Data from 2 placebo-controlled trials of intravenous immunoglobulin (IvIg) conducted in Sweden[28] and Denmark,[29] respectively, may clarify the issue. Patients with 3 or more miscarriages who met the inclusion criteria for participation in the trials were encouraged to conceive and call the clinics as soon as they got pregnant. When the Swedish trial (cohort 1) was concluded after 4 years, 50.6% of the patients were classified as not having achieved pregnancy, only 3.4% had pre-embryonic losses, and 10.1% had embryonic losses (**Fig. 1**). When the Danish trial (cohort 2) was concluded after 6 years, only 14.7% did not report pregnancy, 22.4% had pre-embryonic losses, and 27.9% had embryonic losses. When live birth rates in cohorts 1 and 2 are calculated with the number of recognized pregnancies in the denominator, they become 72.7% and 44.8%, respectively ($P<.005$). The most striking difference between the 2 cohorts is the very much higher frequency of nonconception in cohort 1 than in cohort 2. This may be because of the simple fact that the Danish patients were told that they could not be included in the trial if they contacted the clinic more than

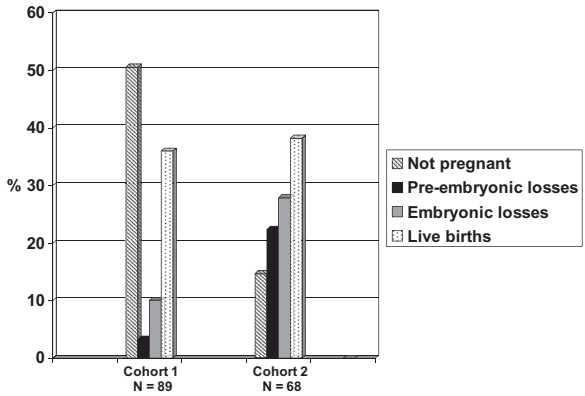

Fig. 1. Classification of pregnancy outcome in 4 categories of patients with RPL eligible for participation in a Swedish (cohort 1) and a Danish (cohort 2) placebo-controlled trial of IvIg. In each cohort, patients receiving IvIg and placebo are combined.

5 days after the missed menstrual period, whereas in the Swedish trial, patients were included only when fetal heart action could be demonstrated in GW 6 to 7. In cohort 2, every conception was probably registered, whereas in cohort 1, many pre-embryonic losses may have been misclassified as nonconception. Another published cohort of patients with RPL that estimated the chance of live birth in the first subsequent pregnancy, also found that 30% of the patients reported no further pregnancies after referral to the clinic or were lost to follow-up.[30] This is again a much higher frequency that the 14.7% in the Danish cohort. As in the Swedish cohort, this difference is probably caused by incomplete follow-up resulting in misclassification of preclinical pregnancy losses as nonconception. When results from different cohort studies are compared, this misclassification becomes nonrandom (see **Table 1**).

Cohorts Estimating Live Births Per Time Unit

The aforementioned examples illustrate the unreliability of studies of RPL cohorts based on the patients' self-reporting of outcome of the first subsequent pregnancy. My group has therefore proposed a more robust method of estimating the prognosis for the patients. We obtained information about subsequent live birth from the national birth registry (which registers 100% of all births in Denmark) for all patients with RPL referred to our clinic from 1986 to 2008. We could then calculate age-specific and prior number of miscarriages–specific proportions of patients who had achieved a live birth within 5 and 15 years after the first consultation (66.7% and 71.1%, respectively).[11] Such an estimate with the rate of live birth in the numerator and the observation time in the denominator is a robust and clinical relevant way to estimate the prognosis, and it is not sensitive to the intensity of monitoring as in the previously discussed methods. However, it can, of course, be used only in countries with a complete and valid national birth registration system. Our estimates of the long-term live birth rates in subsets of patients with RPL, most of whom had attempted pregnancy several times after referral, are generally lower than those reported in the first pregnancy in a frequently quoted cohort,[30] and because our cohort is not sensitive to misclassification, the other cohort probably provides too optimistic rates.

TREATMENT TRIALS

In a time of embracement of evidence-based medicine, all treatments, new as well as established, should undergo testing in clinical trials. With regard to RPL, clinical treatment trials compare miscarriage rates in patients with RPL given one specific treatment in the present pregnancy compared with outcome in a control group not receiving the treatment.

Historical Controls

In some treatment trials, pregnancy outcome in a group of patients treated in the present was compared with outcome in nontreated controls from the past, so-called historical controls.

In a small study of PGS in the treatment of RPL in women aged 35 years,[31] the live birth rate after PGS was for each patient compared with the expected rate in women with similar age and number of miscarriages calculated from a historical cohort.[30] The investigators found that the live birth rate after PGS was higher than the expected rate using the data from the historical cohort. Comparing outcome in contemporary RPL patient groups with historical control groups carries a substantial risk of producing unreliable results (see **Table 1**). It is extremely difficult to be sure that a group of patients with RPL identified in the past is comparable with a present group, as ascertainment of

patients, screening methods, and pregnancy-monitoring procedures have changed substantially with time.

A variant of the use of historical controls is the use of the treated patients as their own controls. In RPL research, the live birth rate in patients with RPL before a specific intervention is compared with the birth rate after the intervention. The use of this method is still seen in publications especially concerning surgical treatment of RPL, but also regarding other treatment modalities, such as metformin.[32] Although using the patients as their own controls at a first glance would appear ideal, this method is severely flawed due to the phenomenon *regression to the mean*.[33] The population of patients with RPL may comprise the following:

- A subset of women who have experienced the miscarriages merely because of chance (repeated embryonic aneuploidies) with a low intrinsic risk of miscarriage in each pregnancy but the women have been sampled asymmetrically from a binominal distribution.
- A subset of women with maternal risk factors with a higher intrinsic risk of miscarriage.

In the former subset, on any subsequent measure (new pregnancy) the mean risk of miscarriage will be closer to the (low) mean of the original population in spite of no intervention due to *regression to the mean*. If they are selected to a treatment trial, the chance of success in the next pregnancy will be excellent in spite of no intervention. In a trial of cerclage, the posttreatment birth rate was statistically significantly higher than before cerclage and it was concluded that cerclage increases the live birth rate. However, using this method, all types of interventions can be proved to be efficient in the treatment of RPL, even placebo.[34] In an RCT of infusion of allogeneic lymphocytes versus placebo in RPL, the patients who received placebo had a live birth rate of 12.5% before they entered the trial and subsequently 47.7% gave birth after placebo infusions. By comparing birth rates before and after placebo by χ^2 test, the statistics claim that placebo is highly efficient in treating RPL. This apparent, but false, improvement of outcome can be completely attributed to the effect of *regression to the mean*.

Randomized Controlled Trials

Proper evaluation of treatments in RPL can be done only in RCTs. In RCTs, patients are prospectively allocated to treatment or no-treatment or placebo by some randomization procedure that cannot be influenced by the researcher. This will increase the chance that confounding variables, which can influence the prognosis, are equally distributed in the groups.

Blinding

In the ideal RCT, allocation to active treatment or a placebo that cannot be distinguished from each other is undertaken according to a randomization list so that the trial is blinded (masked) for both the researchers and patients, a so-called double-blind design. The effects of blinding in RCTs are as follows:

- Blinding of the doctors will ensure the same monitoring and concomitant therapy in both groups
- Blinding of the patients will provide them with the full placebo effect that due to neuroendocrine pathways may exhibit a positive effect on pregnancy outcome in RPL[35]
- Blinding of the patients will diminish the dropout rate and associated poor registration of subsequent outcome in those allocated to the expected less-efficient treatment

Several RCTs in the RPL area are unfortunately not blinded. None of the trials testing the effect of heparin plus low-dose aspirin in patients with RPL with or without APL has compared this intervention against a placebo for heparin but only against low-dose aspirin alone or a peroral placebo for aspirin.[36–38] This is probably due to the reluctance to use a placebo drug for daily subcutaneous injections during the whole period of pregnancy. Although the outcome in RPL, a new miscarriage, can be determined by objective methods that are not impacted by blinding/nonblinding, nonblinding can negatively or positively influence outcome (see **Table 1**) in the groups receiving nontreatment or expected less-efficient treatment compared with the expected more-efficient treatment (tablets vs injections plus tablets).

Although a double-blinded randomized placebo-controlled trial is the best design to obtain valid data concerning treatment effect, this method does not guarantee that the results are trustworthy, as other circumstances can confound or weaken the results.

Confounding in RCTs

Most RCTs in RPL have included a small number of patients and this increases the risk that a confounding variable, such as number of prior miscarriages and age, by chance is unevenly distributed between the 2 allocation groups. In a small RCT, it is not sufficient to show that the mean number of previous miscarriages or age is not statistically significantly different in the 2 allocation groups; these variables can still significantly confound the results. The best way to counteract uneven distribution of prognostic variables in RCTs is by doing block-randomization, whereby separate allocation of patients is done according to groups with comparable number of miscarriages and age.

Inclusion after GW 6

In numerous RCTs, patients with RPL have been allocated and included in the trial only when an ultrasonographic examination has demonstrated a viable intrauterine pregnancy in GW 6 to 7.[28,38,39] This design can unfortunately substantially diminish any effect of the active intervention (see **Table 1**). Twenty-two percent of closely monitored patients with RPL have pre-embryonic losses (see **Fig. 1**), which comprise almost half of all their pregnancy losses.[29] When treatment in RCTs starts only after the demonstration of fetal heart action, almost half of the risk period for miscarriage has thus passed and the spontaneous prognosis at that time is good.[40] Therefore, very large numbers of participants are needed to show any treatment effect. Furthermore, when treatment is started only in GW 6 to 7, the pregnancy has been exposed to the full harmful effect of thrombophilic or immunologic factors for 2 to 3 weeks before initiation of the potential beneficial therapy. The trophoblast or fetus, although viable at that time, may have suffered irreversible damage that cannot be counteracted by initiating active therapy.

Premature discontinuation of an RCT

In many RCTs, at least 1 interim analysis is performed during the conduct of the trial and it is prematurely stopped if the results at this analysis show that a statistically significant effect of the intervention could not be obtained even if the trial was continued until the originally planned number of participants was reached. Many RCTs in the area of RPL have been stopped prematurely after performance of an interim analysis,[28,41,42] and this poses a substantial problem in interpreting the results, especially when they are included in systematic reviews and meta-analyses. There is a great risk that the difference between the outcome in the intervention and nonintervention/placebo groups at the stopping point after the interim analysis due to a random fluctuation is smaller than if the trial had been continued to the planned number of

participants; if the difference had been larger the trial had been allowed to continue. RCTs stopped after interim analyses will therefore be expected to report intervention effects that are smaller than the true effects (see **Table 1**) and, if included in systematic reviews, their scientific quality should accordingly be down-graded.

Inclusion of several outcomes from the same patient

In case series and RCTs in RPL research, several pregnancy outcomes from the same patients are sometimes included.[43,44] This is a methodological flaw, as the outcomes of pregnancies in individual patients with RPL due to the importance of maternal risk factors (eg, age, number of prior losses) are linked variables. The commonly used statistical methods, such as χ^2 test and Fisher test, require that the tested variables are independent and cannot be used. Therefore, only one pregnancy outcome from each patient should be included in trials testing interventions, whether randomized or not.

Systematic Reviews

A systematic review is a review that uses systematic and explicit methods to critically appraise a research topic; statistical methods such as meta-analyses may be used to analyze and summarize the results of the included studies. Both case-control studies and RCTs can be subject to such reviews and their results are considered the highest level of evidence. In RPL research, systematic reviews have in particular focused on intervention/treatment studies. The pooled OR calculated from the combination of results in the meta-analysis is considered a good measure for the overall effect of the intervention/treatment under study. However, because most treatments exhibit different effects in different subsets of patients (eg, men and women or patients with severe vs less severe disease), an important use of meta-analyses is to identify subgroups of patients who have the largest benefit of treatment.[45] I focus my discussion regarding systematic reviews in RPL on those dealing with immunotherapy with IvIg, because these reviews have included the largest numbers of studies and patients and the discussion illustrates the pitfalls and methodological flaws that characterize systematic reviews in the area of RPL treatment. Four different systematic reviews have been published concerning IvIg treatment in RPL,[4,42,46,47] which have provided very different results and conclusions. This variability is in my view caused by either (1) a failure to recognize the need of doing separate analysis in relevant subgroups of patients or (2) by unfounded exclusions of RCTs in the systematic reviews. A Cochrane review[4] did not find any difference in live birth rate between IvIg and placebo in all included patients with RPL but made no distinction between patients with primary and secondary RPL, which would be relevant, as 2 published trials had found the IvIg efficient exclusively in the latter group.[29,48] In contrast, another systematic review based on almost the same patients recognized the relevance of doing this subgroup analysis and found the live birth rate significantly higher in IvIg-treated than in placebo-treated patients with secondary RPL.[46]

The 2 most recent systematic reviews on IvIg in RPL included only patients with "unexplained" RPL, which meant that trials that had included patients with RPL positive for APL according to the assays used in the individual clinics and according to the local cutoff values were excluded.[42,47] The investigators considered that 2 RCTs included APL-positive patients and all 92 patients in these RCTs were completely excluded from the systematic reviews comprising 25.3% of all 364 patients with RPL participating in RCTs of IvIg. Exclusion of this substantial proportion of all randomized patients with RPL appears to be an enormous waste of valid information. In one of the excluded RCTs, none of the patients in fact had APL and in the other almost all patients who tested APL-positive had very low titers.[49] Only 3% would be

considered APL-positive according to the current criteria.[50] On the other hand, an RCT that was not excluded did in fact include APL-positive patients[51] and several of the other RCTs did not report details of APL assays or cutoff levels. Because both excluded RCTs had found a substantial effect of IvIg in patients with secondary RPL, doing the meta-analyses without these resulted in the conclusion in both systematic reviews that there is no benefit of IvIg either in primary or secondary RPL.[42,47] As stated by my group and others, we find the methodology of these systematic reviews questionable and without the exclusions it can be proved that IvIg is indeed efficient in secondary RPL.[49,52] As previously mentioned, it is legal and desirable to do meta-analyses in subgroups of patients to identify subgroups that benefit best from treatment, and such a subgroup could be patients with RPL without APL. However, a systematic review excluding these patients should of course exclude only those patients who are positive for APL and not complete RCTs with only a tiny minority of participants being APL-positive. Furthermore, in most published relevant RCTs, insufficient information is given about the APL assays and cutoff values, which may vary substantially between the trials. To carry out a meta-analysis of therapies in APL-negative patients with RPL based on good research methodology, it is therefore necessary to collect data on APL levels for each included patient. In most cases, these data on individual patients can be obtained only by contacting the investigators of the RCTs. If such a collection of raw data is not done, the resulting systematic review and meta-analysis will produce biased results.

This discussion illustrates that even conclusions from systematic reviews should be evaluated critically and a systematic review is not per se the highest level of evidence. If they are not conducted according to rigorous and systematic rules, they will end up being no more evidence-based than the narrative reviews of the old days (see **Table 1**).

SUMMARY

Most causes of RPL are poorly elucidated and may have a more multifactorial etiology than infertility and in many instances the cause-effect relationship is unclear. Furthermore, the RPL population is much smaller that the infertile population, thereby making it difficult to conduct studies with adequate statistical power.

These problems render research in RPL inherently difficult and only methodologically high-quality studies are expected to produce useful results. Unfortunately, such high-quality research studies are rare in the area of RPL.

All types of studies in RPL research are characterized by methodological pitfalls, which are often not recognized by the researchers or readers. Failure to recognize these can completely invalidate studies in the area. It is hoped that this review, by setting focus on the problem, can help improve studies in the area of RPL in the future.

REFERENCES

1. The Practice Committee of the American Society for Reproductive Medicine. Evaluation and treatment of recurrent pregnancy loss: a committee opinion. Fertil Steril 2012;98:1103–11.
2. Royal College of Obstetricians and Gynaecologists. The investigation and treatment of couples with recurrent first-trimester and second-trimester miscarriage. RCOG Green-top Guideline No. 17. 2011.
3. Jauniaux E, Farquharson RG, Christiansen OB, et al. Evidence-based guidelines for the investigation and medical treatment of recurrent miscarriage. Hum Reprod 2006;21:2216–22.

4. Porter TF, LaCoursiere Y, Scott JR. Immunotherapy for recurrent miscarriage. Cochrane Database Syst Rev 2006;(2):CD000112. http://dx.doi.org/10.1002/14651858.
5. The Practice Committee of the Society for Assisted Reproductive Technology, Practice committee of the American Society for Reproductive Medicine. Preimplantation genetic testing: a Practice Committee opinion. Fertil Steril 2008;90:S136–46.
6. Nybo Andersen AM, Wohlfart J, Christens P, et al. Maternal age and fetal loss: population based register linked study. BMJ 2000;320:1708–12.
7. van den Boogaard E, Cohn DM, Korevaard JC, et al. Number and sequence of preceding miscarriages and maternal age for the prediction of antiphospholipid syndrome in women with recurrent miscarriage. Fertil Steril 2013;99:188–92.
8. Jaslow CR, Careney JL, Kutteh WH. Diagnostic factors identified in 1020 women with two versus three or more recurrent pregnancy losses. Fertil Steril 2010;193:1234–43.
9. Wilcox AJ, Weinberg CR, O'Connor JF, et al. Incidence of early loss of pregnancy. N Engl J Med 1988;319:189–94.
10. Kolte AM, van Oppenraaij RH, Quenby S, et al. Biochemical pregnancy loss and pregnancy of unknown location are prognostically important for unexplained recurrent miscarriage. Under review.
11. Lund M, Kamper-Jørgensen M, Nielsen HS, et al. Prognosis for live birth in women with recurrent miscarriage. What is the best measure of success? Obstet Gynecol 2012;119:37–43.
12. Axelsson G, Rylander R. Validation of questionnaire reported miscarriage, malformation and birth weight. Int J Epidemiol 1984;13:94–8.
13. Wilcox AJ, Horney LF. Accuracy of spontaneous abortion recall. Am J Epidemiol 1984;120:727–33.
14. Out HJ, Bruinse HW, Christiaens GC, et al. A prospective, controlled multicenter study on the obstetric risks of pregnant women with antiphospholipid antibodies. Am J Obstet Gynecol 1992;167:26–32.
15. Regan L, Braude PR, Hill DP. A prospective study of the incidence, time of appearance and significance of anti-paternal lymphocytotoxic antibodies in human pregnancy. Hum Reprod 1991;6:294–8.
16. Lissauer D, Piper K, Goodyear O, et al. Fetal-specific CD8+ cytotoxic T cell responses develop during normal human pregnancy and exhibit broad functional capacity. J Immunol 2012;189:1072–80.
17. Wegienka G, Havstad S, Bobbitt KR, et al. Within-woman change in regulatory T cells from pregnancy to the postpartum period. J Reprod Immunol 2011;88:58–65.
18. Gamill HS, Guthrie KA, Aydelotte TM, et al. Effect of parity on fetal and maternal microchimerism: interaction of grafts within a host? Blood 2010;116:2706–12.
19. O'Donoghue K, Chan J, de la Fuente J, et al. Microchimerism in female bone marrow and bone decades after fetal mesenchymal stem-cell trafficking in pregnancy. Lancet 2004;364:179–82.
20. Raghupathy R, Al Mutawa E, Makhseed M, et al. Modulation of cytokine production by dydrogesterone in lymphocytes from women with recurrent miscarriage. BJOG 2005;112:1096–101.
21. von Wolff M, Thaler CJ, Strowitzki T, et al. Regulated expression of cytokines in human endometrium throughout the menstrual cycle: dysregulation in habitual abortion. Mol Hum Reprod 2000;6:627–34.

22. Schumacher A, Brachwitz N, Sohr S, et al. Human chorionic gonadotrophin attracts regulatory T cells into the fetal-maternal interface during human pregnancy. J Immunol 2009;182:5488–97.
23. Piosek ZM, Goegebeur Y, Klitlou L, et al. Plasma TNF-α levels are higher in early pregnancy in patients with secondary compared with primary recurrent miscarriage. Am J Reprod Immunol 2013;70:347–58.
24. Calleja-Agius J, Jauniaux E, Pizzey AR, et al. Investigation of systemic inflammatory response in first trimester pregnancy failure. Hum Reprod 2012;27:349–57.
25. Quack KC, Vassiliadou N, Pudney J, et al. Leukocyte activation in the decidua of chromosomally normal and abnormal fetuses from women with recurrent abortion. Hum Reprod 2001;16:949–55.
26. Bradford-Hill A. The environment and disease: association or causation? Proc R Soc Med 1965;9:295–300.
27. Lund M, Nielsen HS, Hviid TV, et al. Hereditary thrombophilia and recurrent pregnancy loss: a cohort study of pregnancy outcome and obstetric complications. Hum Reprod 2010;25:2978–84.
28. Jablonowska B, Selbing A, Palfi M, et al. Prevention of recurrent spontaneous abortion by intravenous immunoglobulin: a double-blind placebo-controlled trial. Hum Reprod 1999;14:838–41.
29. Christiansen OB, Pedersen B, Rosgaard A, et al. A randomized, double-blind, placebo-controlled trial of intravenous immunoglobulin in the prevention of recurrent miscarriage: evidence for a therapeutic effect in women with secondary recurrent miscarriage. Hum Reprod 2002;17:809–16.
30. Brigham SA, Conlon C, Farquharson RG. A longitudinal study of pregnancy outcome following idiopathic recurrent miscarriage. Hum Reprod 1999;14:2868–71.
31. Munné S, Chen S, Fischer J, et al. Preimplantation genetic diagnosis reduces pregnancy loss in women aged 35 years and older with a history of recurrent miscarriage. Fertil Steril 2005;84:331–5.
32. Glueck CJ, Wang P, Goldenberg N, et al. Pregnancy outcomes among women with polycystic ovary syndrome treated with metformin. Hum Reprod 2002;17:2858–64.
33. Yudkin PL, Stratton IM. How to deal with regression to the mean in intervention studies. Lancet 1996;347:241–3.
34. Christiansen OB. Transabdominal cervicoisthmic cerclage in the management of recurrent second trimester miscarriage and preterm delivery. BJOG 1996;103:595–6.
35. Stray-Pedersen B, Stray-Pedersen S. Etiological factors and subsequent reproductive performance in 195 couples with a history of habitual abortion. Am J Obstet Gynecol 1984;148:140–6.
36. Kutteh WH. Antiphospholipid antibody-associated recurrent pregnancy loss: treatment with heparin and low-dose aspirin is superior to low-dose aspirin alone. Am J Obstet Gynecol 1996;174:1584–9.
37. Rai R, Cohen H, Dave M, et al. Randomised controlled trial of aspirin and aspirin plus heparin in pregnant women with recurrent miscarriage associated with phospholipid antibodies (or antiphospholipid antibodies). BMJ 1997;314:253–7.
38. Kaandorp SP, Goddijn M, van der Post JAM. Aspirin plus heparin or aspirin alone in women with recurrent miscarriage. N Engl J Med 2010;362:1586–96.
39. The German RSA/IVIV Group. Intravenous immunoglobulin in the prevention of recurrent miscarriage. BJOG 1994;101:1072–7.

40. Stern J, Coulam CB. Mechanism of recurrent spontaneous abortion. I. Ultrasonographic findings. Am J Obstet Gynecol 1992;166:1844–52.
41. Ober C, Karrison T, Odem RR, et al. Mononuclear-cell immunisation in prevention of recurrent miscarriages: a randomised trial. Lancet 1999;354:365–9.
42. Stephenson MD, Kutteh WH, Purkiss S, et al. Intravenous immunoglobulin and idiopathic secondary recurrent miscarriage: a multicentered randomized placebo-controlled trial. Hum Reprod 2010;25:2203–9.
43. Brenner B, Hoffman R, Blumenfeld Z, et al. Gestational outcome in thrombophilic women with recurrent pregnancy loss treated with enoxaparin. Thromb Haemost 2000;83:693–7.
44. Grandone E, Brancaccio V, Colaizzo D, et al. Preventing adverse obstetric outcomes in women with genetic thrombophilia. Fertil Steril 2002;78:371–5.
45. Thompson SG, Higgins JPT. Can meta-analysis help target interventions at individuals most likely to benefit? Lancet 2005;365:341–6.
46. Hutton B, Sharma R, Fergusson D, et al. Use of intravenous immunoglobulin for treatment of recurrent miscarriage: a systematic review. BJOG 2007;114:134–42.
47. Ata B, Tan SL, Shehata F, et al. A systematic review of intravenous immunoglobulin for treatment of unexplained recurrent miscarriage. Fertil Steril 2011;95:1080–5.
48. Christiansen OB, Mathiesen O, Husth M. Placebo-controlled trial of treatment of unexplained secondary recurrent spontaneous abortions and recurrent late spontaneous abortions with i.v. immunoglobulin. Hum Reprod 1995;10:2690–5.
49. Christiansen OB, Larsen EC, Husth M, et al. Intravenous immunoglobulin and recurrent miscarriage. Fertil Steril 2011;7:e35.
50. Myakis S, Lockshin MD, Atsuma T, et al. International consensus statement on an update on the update of the classification criteria for the definite antiphospholipid syndrome (APS). J Thromb Haemost 2006;4:295–306.
51. Mueller-Eckhardt G, Mallmann P, Neppert J, et al. Immunogenetic and serological investigations in nonpregnant and in pregnant women with a history of recurrent spontaneous abortions. J Reprod Immunol 1994;27:95–109.
52. Clark DA. Intravenous immunoglobulin and idiopathic secondary recurrent miscarriage: methodological problems. Hum Reprod 2011;26:2586–91.