Unexplained sporadic and recurrent miscarriage in the new millennium: a critical analysis of immune mechanisms and treatments

David A. Clark¹,⁵, Carolyn B. Coulam², Salim Daya³ and Gerard Chaouat⁴

¹Departments of Medicine, Molecular Medicine and Pathology, Obstetrics and Gynecology, McMaster University, Rm. 3V39, Hamilton, Ontario, Canada L8N 3Z5, ²Sher Center for Reproductive Medicine, Chicago, USA and Chair, RMITG Study Group, ³Departments of Obstetrics and Gynecology, and Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada and ⁴Maitre de Recherche, INSERM U 131, Hopital Antione Beclere, Clamart, France

There have been important advances in basic science investigation of mechanisms underlying spontaneous miscarriages which lend support to empirical treatments such as intravenous immunoglobulin G and allogeneic leukocyte immunotherapy. The results from clinical trials of these and other proposed treatments have been problematic. There is only one published meta-analysis of sufficient power and appropriate stratification to qualify as Level 1 evidence, and that deals only with leukocyte immunotherapy. Here we critically review current trials and their flaws, update the meta-analysis, and comment on potential new approaches. Inadequate sample size, better definition of heterogeneity, and proper stratification to minimize the effects of heterogeneity remain as problems. Verification that the experimental or test treatment was active in producing the expected alteration in immunophysiology in the recipient is lacking in most trials; use of stored rather than fresh allogeneic leukocytes appears problematic. Hidden biases that affect trial significance emerge with critical analysis, and the focus on apparent ‘high quality’ of design in published reports may be misleading. We conclude that there seem to be enough patients to conduct clinical trials of sufficient size to achieve adequate power to test therapies showing promise in pilot studies, but at present, the only Level 1 evidence concerns leukocyte immunotherapy which appears to increase the chance of a live birth if given to appropriate patients.

Keywords: cytokines/immunotherapy/IVIg/meta-analysis/miscarriage

TABLE OF CONTENTS

Introduction
Immunological mechanisms in pregnancy failure
Interventions to prevent pregnancy loss
Conclusions
Acknowledgements
References

Introduction

Although many advances in reproductive medicine have been made during the past quarter of a century, miscarriages remain the most common complication of pregnancy. Miscarriages affect 15% of women, primarily in the first trimester (Daya, 2000). Whilst most are sporadic and non-recurrent, there is a subset comprising 2–5% of couples that suffers recurrent miscarriage (Mills et al., 1988). These repetitive losses suggest the presence of a specific cause, and much work has been done to try to identify such underlying mechanisms. Women whose first spontaneous abortion is karyotypically normal have an increased risk of a recurrence (Boue et al., 1973). By identifying causes of, and remedies for, recurrent loss of ‘normal’ embryos, one hopes to prevent loss of all karyotype normal embryos. Examples of non-recurrent losses which might be saved included the 5% of first losses which affect otherwise normal embryos, one hopes to prevent loss of all karyotype normal embryos. Examples of non-recurrent losses which might be saved included the 5% of first losses which affect otherwise normal embryos (one-third of all miscarriages affecting women), and the 30–40% of women undergoing IVF, who achieve pregnancy but abort (>80% of transfers of three preimplantation embryos should include at least one of normal karyotype).

Recurrent spontaneous abortion (RSA) can be partner-specific or non-specific (Coulam et al., 1976). The latter group is not frequently identified, if only because few women do the ‘experiment’. Some recurrent loss problems only appear when a woman takes a new partner. Most recurrent miscarriage problems involve a series of consecutive losses. Those with a single
liveborn (L), usually the first followed by a series of miscarriages (A) (e.g. LAAA [equiv]), are considered to have a similar problem to those with only miscarriages (i.e. AAA [equiv]) (Mowbray and Underwood, 1991). A mixture of liveborn and miscarriages (e.g. LLAALALA [equiv]) is considered to probably have genetic causes (i.e. chromosome abnormalities) if only because of their more favourable prognosis compared to AAAA [equiv] (Warburton and Fraser, 1964). That such patterns are thought to have clinical implications is one indicator of the problem of heterogeneity. As we will argue, when one is thinking about treatment, far more information is needed than merely obstetric history. A variety of investigations is usually carried out to try to identify some abnormality upon which one can blame the recurrent miscarriages. These include uterine anomaly (e.g. septum), endocrine disturbance including ‘luteal phase defect’, antiphospholipid antibodies, endometritis and parental chromosome abnormality (Stray-Pedersen and Stray-Pedersen, 1984). Endometriosis may also represent a risk, albeit controversial (Daya, 1996). Whether such ‘causes’ are valid has been questioned. For example, with a parental chromosome abnormality, 50% of conceptions should abort, but the actual loss rate is only 15% (P < 0.05) (Clark et al., 1996). Further, the list of ‘explanations’ fails to take into consideration the problem of recurrent chromosomally abnormal embryos which have been designated as ‘unexplained’ because they were not karyotyped. The following paragraphs summarize current understanding of mechanisms of loss and effectiveness of current treatment strategies for recurrent ‘unexplained’ miscarriage of normal karyotype embryos. We will mention, where appropriate, newer ‘causes’ proposed for such losses, e.g. hyperhomocysteinaemia, thrombophilia, and antiphospholipid antibody syndrome/autoimmunity, but will not review these mechanisms or possible treatments in detail.

**Immunological mechanisms in pregnancy failure**

Karyotype analysis of products of conception suggests that chromosomal aberrations in the spermatozoon, oocyte, or early zygote can explain about two-thirds of sporadic pregnancy losses (Boeuf et al., 1973; Hassold et al., 1980; Coulam et al., 1996). If such ‘accidents’ explained recurrent loss, the probability of three or more spontaneous abortions in a row resulting from ‘accidents’ would account for ≈5% of the observed incidence of losses. On this basis, calculations turned to the role of immunology and possible rejection of the paternal alloantigen-bearing embryo as an ‘allograft’. Allogeneic blood transfusion, which may prolong survival of kidney allografts (Bordin et al., 1994), was tried, and investigators searched for elusive ‘blocking antibodies’ that could prolong allograft survival in animal models (Check et al., 1997). Fortunately, it was possible to access early pregnancy failures and ascertain the karyotype of the trophoblast in >94% of cases (Coulam et al., 1996), and fortunately, an animal model of spontaneous early pregnancy failure which was partner-specific was discovered (Clark et al., 1999a). The latter has provided useful clues to investigations that might be profitable in humans. Karyotypic analysis of >200 abortuses from women with ‘unexplained’ primary recurrent abortion with the same partner showed that 55% of losses were abnormal, and 35% of those with secondary recurrent loss (one live birth) were also chromosomally abnormal, in agreement with data published by others (Coulam et al., 1996). Analysis by number of prior miscarriages indicated that there is a subset of couples with a 60–70% risk of a recurrence, not a 10% risk (23 × 15%) as had been assumed (Coulam et al., 1996). Although repeated attempts at conception should eventually lead to a normal embryo in such couples, it is possible that some have low risks of recurrence and others a risk near 100%. Indeed, trisomic pregnancies are noted for the high risk of a repeat trisomy (Naylor and Warburton, 1979). As there is currently no way, other than by karyotyping, to identify those with losses due to abnormal embryos, we have advocated a change in practice such that all losses after a first miscarriage should be karyotyped (Coulam et al., 1996; Clark and Coulam, 1996). This approach requires close monitoring by β-human chorionic gonadotrophin (β-HCG) concentrations and ultrasound in the next pregnancy with early chorionic villous biopsy as soon as the pregnancy begins to fail. A second reason for collecting this information is that such patients may inappropriately end up in an allogeneic leukocyte immunotherapy trial after their third loss; if the second and third miscarriages are chromosomally abnormal, they should be excluded. There is a third reason for seeking karyotype data. If one critically examines the rate of success of the next pregnancy in cohorts of patients who are not given treatment other than tender loving care (TLC), the rate of success ranges from 50 to 85% (Clark et al., 1997). Some of the variation can be explained by the worsening prognosis when the number of prior losses of more than three increases (Clark and Coulam, 1996; Coulam et al., 1996), but the majority of the patients have had only three unexplained losses. We suspect that the remarkable variation in success rates, apart from simple statistical variability, may reflect a variation in the incidence of chromosome abnormalities in the embryos. Those reporting 80% success rates might be less confident of the effectiveness of their intervention if they knew that the frequency of chromosome errors in their patients, embryos was 5% and not 55%. For example, if the failure rate in the next pregnancy after three unexplained miscarriages and no live births were 60%, and 55% of the losses are due to chromosomal abnormalities, then 33% (55% of 60%) are lost due to lethal chromosome abnormalities; the maximum possible success rate is 67%, so an 80% success rate is not possible. On the other hand, if the population has a 10% loss rate due to chromosomal abnormalities, then a 90% success rate should be achievable, and 80% is less than might be achieved with effective therapy. Not infrequently in such studies, ‘treatment’ was begun after the usual time of onset of the process leading to miscarriage (Clark and Daya, 1991). The frequency of karyotype abnormal recurrent abortuses could provide valuable clues in searching for causes of recurrent chromosome abnormalities, e.g. putative environmental toxins (Warburton, 1985).

It has been suggested that certain types of early pregnancy failure identified by ultrasound might be helpful in estimating the likelihood of an ‘uncorrectable’ chromosome abnormality, i.e. blighted ovum as distinct from missed abortion where the fetal heartbeat is lost (FHL, also called intrauterine fetal demise, IUFD). Such surrogate information could partially solve the problem created by lack of fetal trophoblast karyotype data for most patients who are referred for immunological assessment. Trophoblast of abnormal karyotype represents abnormal cells which are thought to divide more slowly, thereby generating a
subnormal rate of rise of βHCG that is detectable prior to the sixth week of gestation. These embryos are not thought to be 'rescuable' with immunological therapies. A high frequency of failure to establish fetus–placenta vascular connection in blighted ova has been shown (Meegdes et al., 1985)—a problem that altering maternal immune status cannot correct. Whilst the frequency of lethal karyotype abnormalities may be higher in blighted ova compared to FHL/IUF-type miscarriages, the association is imperfect (Sorokin et al., 1991; Minelli et al., 1993).

The frequency of losses occurring prior to the sixth week of gestation and before the woman may realize she is pregnant (i.e. occult losses) is calculated to be two to five times greater than the clinical miscarriage rate, but the estimated frequency of karyotype anomalies at the occult loss stage is unable to account for all of these losses (Lea and Clark, 1991). Therefore, identification of karyotype is necessary for any study of mechanism or loss; implications of early occult loss and blighted ovum versus FHL/IUFD patterns of loss for immunological therapies will be discussed below.

There are data from the laboratory mouse, which now appear relevant to those aborting chromosomally normal embryos. In this species, the frequency of chromosome abnormalities is extremely low, <6% (de Boer et al., 1991). Mating of CBA/J female mice with allogeneic DBA/2 males generates pregnancies susceptible to a high rate of spontaneous abortion; this is partner specific since changing either the female or male strain leads to pregnancies with abortion rates of <10% (Clark et al., 1999a). The spontaneous abortion rate in matings of CBA/J females to DBA/2 males can range from 10 to 50%, depending on the level of 'stress' and endogenous bacterial flora which serve as eliciting triggers for loss. Both activation local production in the uterine lining cells of the proinflammatory cytokines tumour necrosis factor (TNF)-α and interferon (IFN)-γ which are essential for abortions to occur (Clark et al., 1999a). Macrophages (a source of proinflammatory cytokines such as TNF-α, interleukin (IL)-12 and IL-1 and natural killer (NK) cells (a source of IFN-γ and some TNF-α) infiltrate implantation sites destined to abort, and abortions begin at the time of formation of a vascularized placenta, about 5 days after implantation (corresponding to ~6 weeks of gestational age of a human embryo) (Clark, 1991a). It was thought that NK cells and/or macrophages might be killing trophoblasts, because mouse placental trophoblasts proved to be resistant to mechanisms causing allograft rejection (Clark et al., 1999a). However, only NK cells activated in the presence of IL-2 appear able to lyse placental trophoblast, and IL-2 has not been found in the decidua of aborting mice. More recently, it has been shown that abortions are probably caused by TNF-α plus IFN-γ activating production of the novel pro thrombinase fgl2 in trophoblast and in decidua (Clark et al., 1998). This procoagulant leads to deposition of fibrin and activation of polymorphonuclear leukocytes that can destroy the vascular supply to the placenta. This discovery was made during experiments done to test the requirement for NK cells and/or macrophages as 'effector cells' causing the demise of embryos in abortion-prone CBA/J×DBA/2 matings 1 a model in which both cytokines are important. On one hand, it had been suggested that NK cell-derived IFN-γ activated macrophages produce toxic levels of nitric oxide, but on the other hand, macrophage-derived TNF-α might convert NK cells into lymphokine-activated killer cells (LAK) (Clark et al., 1998).

Injection of either TNF-α or IFN-γ into pregnant mice on day 7.5 of gestation normally boosts the rate of abortions that first become apparent on day 9.5 of gestation. Surprisingly, if the females had been depleted of either macrophages or NK cells, neither cytokine had any effect. However, if TNF-α was administered together with IFN-γ on day 7.5, >80% of the embryos of the macrophage-depleted or NK cell-depleted mice were aborted. Studies using genetically modified females lacking an IFN-γ response-1 element mated to normal males did not abort in response to an injection of TNF-α plus IFN-γ; accordingly, we postulated that the cytokines acted on maternal cells (but not NK cells or macrophages) and did not work by a direct effect on fetal trophoblast to cause apoptosis (Clark et al., 1999). Induction by cytokines of vascular endothelial cell procoagulants was suggested, and, in support of this idea, cytokine-triggered abortions were almost completely prevented by injecting the females with purified antibody neutralizing the novel prothrombinase, fgl2 (Clark et al., 1998).

Fgl2 was only one of many potential procoagulants that could have been stimulated by cytokines, and the only murine prothrombinase against which neutralizing antibody was available (Clark et al., 1999a). Fgl2 directly converts prothrombin to thrombin, and thrombin then activates endothelial cells to produce IL-8 (MIP2 in the mouse); IL-8/MIP2 recruits polymorphonuclear leukocytes (PMNL) which are able to kill cytokine-activated endothelium (Clark et al., 1998). The importance of generation of thrombin was established by treating the pregnant females with heparin (which activates antithrombin III) or hirudin (a direct antithrombin); both reduced the abortion rate. The importance of PMNL in abortions was shown by injecting monoclonal anti-PMNL antibody (Clark et al., 1998, 1999a). We have recently shown that TNF-α plus IFN-γ enhances fgl2 mRNA expression in both fetal trophoblast and in maternal decidua; there is increased fibrin deposition, and separation of the embryo from the decidua where the fgl2+ trophoblast and decidua meet, in association with an infiltrate of PMNL (Clark et al., 2001a).

From studies in mice, we also have new data suggesting that the processes causing occult pregnancy failure may be distinct from processes causing failure of embryos which have established a vascularized placenta. Neoplastic cells can grow to a size of ~3 mm nourished by diffusion, but, to exceed this size, they must recruit host blood vessels. The cells at the centre of neoplasms that survive by diffusion alone are nevertheless nutritionally compromised, and actively dividing and migrating cells at the core of an embryonic implantation are likely to be compromised at 2–3 mm diameter. In the mouse, this size is reached at approximately days 9–9.5 of gestation when a distinct placenta forms; fetus–trophoblast vascularization and trophoblast–maternal blood contact begins at day 6.5 (Davidson et al., 2000). Recent studies in mice indicate that occult losses are triggered by T cells with αβ-type receptors, particularly NKT cells, the cytolsin perforin may be required, and the C3 component of complement is involved (Arck et al., 1997a; Ito et al., 2000; Xu et al., 2000). It was shown many years ago that transplantation immunity against paternal histocompatibility antigens, primarily minor antigens, could reject blastocysts placed under the kidney capsule, whereas later-appearing trophoblasts, which form the placenta, are unrejectable (Clark, 1991b).

NKT cells are mononuclear leukocytes bearing NK lineage markers and T cell receptors either αβ or γδ. NKT cells are
thought to be members of the antigen-non-specific innate (natural) defence system which have partly evolved towards antigen-specific T cells. NKT cells differ from T cells of the adaptive immune system in having a limited number of antigen patterns that can be recognized and react rapidly to antigen by releasing a pre-programmed set of cytokines.

Blastocysts placed in the uterus of preimmunized mice resist rejection, and protection may be explainable in part by presence of cells in decidua that produce a cryptophanase that inactivates, at short range, maternal effector T cells (Munn et al., 1998). Bacterial endotoxin (lipopolysaccharide) terminates embryo viability at the occult stage (day 4.5–8.5 in the mouse), but this does not require TNF-α (acting via TNF-αR1, P55) (Arck et al., 1997b). By contrast, resorptions (abortions) that begin on or after day 9.5 of gestation do require TNF-αR1 but do not require perforin (Arck et al., 1997b; Stallmach et al., 1998). It is possible that complement may participate in classical abortions, and it is tempting to speculate that activation might be achieved via the serine protease activity of fgl2. Cells preventing cytokine-triggered abortions in mice, by contrast to classical αβ T cells, express γδ T cell receptor (γδTCR). Recognition of antigen by γδ TcR is more akin to antibody–antigen interactions; soluble antigens and antigens unassociated with MHC molecules can be recognized, atypical MHC Class I-b antigens can be recognized, and neither CD4 nor CD8 may be needed to act as a co-receptor, unlike MHC-restricted recognition of peptides bound to the groove of classical MHC determinants on antigen-presenting cells. The importance of γδ T cells will be discussed in more detail below.

Numerous observations made using mice have considerable appeal with respect to providing insight into understanding recurrent human miscarriages. Human placental trophoblast cultured in vitro is likewise resistant to cytotoxic effects of T and NK cells (indeed, in studies with fresh uncultured trophoblast, even lymphokine-activated killer cells were unable to cause lysis) (reviewed in Clark et al., 1999a). Further, a vascular lesion can explain why the first clinically detectable sign of impending miscarriage in humans is usually a plateauing or drop in βHCG concentration; βHCG is released exclusively by syncytiotrophoblast that is in direct contact with maternal blood, and if the blood supply is compromised, less hormone will enter the systemic circulation. In a pilot study using in-situ hybridization for fgl2 mRNA, we have found that increased expression is associated with miscarriage of normal karyotype embryos, by contrast with abnormal karyotype embryos (Clark et al., 1999b, 2001a). Increased expression of fgl2 mRNA is also associated with increased intensity of immunostaining for fgl2 protein in mice and human uterine tissue (Knackstedt et al., 2001). NK cells play an important role in spontaneous abortions in mice, and women with high numbers of circulating NK cells have a higher risk of miscarriage when they eventually undertake a pregnancy (Aoki et al., 1995). Further, in 37–56% of patients with unexplained recurrent abortion, there was a dramatic increase in Th1>Th2 cytokine mRNA in the endometrium not present in any of 10 normal fertile controls (Lim et al., 2000), increased infiltration of decidua in 50% of recurrent miscarriage patients by CD56γδ16+ blood-type NK cells (as distinguished from the novel endometrial/granulated CD56γδ+ cells which lack CD16) (Lachapelle et al., 1996), and in 50–60% of patients, the presence in blood of cells which make an IFN-γ-like material in response to sperm or trophoblast antigen(s) (Hill et al., 1995). The identity of these blood leukocytes remains obscure. Classical T cells (with αβ receptors for antigen) and NK cells do not recognize trophoblast, except possibly via inhibitor receptors (CD94/NKG2) for non-classical MHC Class I-b antigens such as HLA-G and HLA-E (King et al., 1997). However, T cells and NKT cells with γδ receptors for antigen specifically Vγ1 do react to trophoblast and are implicated in mice with abortion of the vascularized placenta (Heyborne et al., 1994; Arck et al., 1997a,b, 1999a,b). The γδ T cell receptor may ‘see’ certain MHC Class I-b antigens, especially if modified by having bound peptides of ≥12 amino acids, bacterial products or fragments of heat shock protein 65 (Heyborne et al., 1994; Clark, 1999a); this is distinct from T cells with TcR αβ which recognize nine amino acid peptides bound in the groove of Class I antigens, and which have not been found able to recognize trophoblast. In the mouse, a large proportion of the NK cells isolated from the decidua of abortion-prone CBA/J×DBA/2 matings have γδ receptors and are NKT cells; eliminating these cells prevents abortions (Arck et al., 1997a, 1999a; Clark et al., 1999a; Clark and Croitoru, 2001). By contrast, Vγ1γδ T cells make cytokines such as IL-10, TGF-αβ, that inhibit NK cell and macrophage activity (Arck et al., 1997a, 1999a; Clark, 1999c). Although cells with γδ TcR appear less predominant in human decidua, both γδ-only and NK-γδβT cells have been described (Mincherva-Nilsson et al., 1997; Clark, 1999; Clark et al., 1999a); γδβT cells are also present in human endometrium (Flynn et al., 2000). However, there has not yet been a systematic investigation to correlate an increased frequency of NK-γδT cells and increased Th1 cytokine production with recurrent miscarriage of chromosomally normal embryos. There is, however, an apparent increase in decidua of women losing chromosomally normal embryos in the ratio of CD16δ δ CD56γδ classical NK cells of the type found in blood to CD16δ CD56γδ atypical decidual NK cells (Yamamoto et al., 1999).

In the mouse, Th2 cytokines such as IL-4 and IL-10 and the Th3 cytokine TGF-αβ have been associated, along with release of a progesterone-induced blocking factor (PIBF), with pregnancy success (Clark et al., 1999a; Szekeres-Bartho et al., 1999). Matings of CBA/J females to BALB/c males (the same MHC as DBA/2) have a low abortion rate and such pregnancies prime the CBA/J female so that when subsequently mated with DBA/2, she has a low abortion rate (Chauvat et al., 1988; Clark et al., 1999a). This can also be achieved by pre-immunizing CBA systemically or into the uterus with BALB/c leukocytes; as few as 1000 cells into the uterus may suffice to induce some degree of protection (Chauvat et al., 1988). Protection is mediated by CD8+ T cells bearing progesterone receptors which produce IL-10 and PIBF (Clark et al., 1999a; Clark, 1999c; Szekeres-Bartho et al., 1999). These same cells have been associated with successful human pregnancy, are Vγ1γδ T cells and hence able to recognize trophoblast (Heyborne et al., 1994; Polgar et al., 1999). It is thought that, once primed to minor paternal antigens in the context of paternal major histocompatibility complex (MHC) Class I-b, there is improved recognition of such antigens expressed on the implanting embryo, and thus suppression of the Th1 response, but protection is short-lived consistent with lack of memory in γδ T cells and a requirement for continuous stimulation to maintain activation (Baines et al., 1996; Clark,
1999a; Hayday, 2000). This suppression is only obligatory where there are high levels of Th1 cytokines. On this basis, it has been suggested that there may be a rationale to immunizing women using their husband’s mononuclear leukocytes (or possibly using third party donor leukocytes) to prevent spontaneous abortions (Clark, 1999a). Although the CBA/I×DBA/2 model requires third party cells, there is also a model (B10×B10.A) of spontaneous abortion where paternal cells are effective (Chaoau et al., 1988). In humans, allogeneic leukocytes may also activate monocytc cells that inhibit NK cells (Higuchi et al., 1995). Further characterization of these cells is required.

Allogeneic leukocyte administration is known to promote a Th1→Th2 cytokine shift (Babcock and Alexander, 1996; Kirkley et al., 1998). Our recent pilot study showing up-regulation of fgl2 prothrombinase, in biopsies of incipiently failing karyotype normal human embryos but not in karyotype abnormal embryos (Clark et al., 1999b; Knackstedt et al., 2001), is consistent with increased clotting in loss of karyotype normal embryos (Salafia et al., 1993). Not all losses associated with increased clotting may be due to unopposed Th1 cytokine up-regulation of fgl2. There is a growing emphasis on abnormal regulation of blood clotting and importance of procoagulants, such as homocysteine, prothrombin mutations, thrombophilia, and antiphospholipid antibodies in human recurrent miscarriage patients (Clark et al., 1999a; Coumans et al., 1999; Blumenfeld and Brenner, 1999; Spina et al., 2000; Younis et al., 2000). A Th1→Th2 cytokine shift is unlikely to protect against all varieties of procoagulant-induced miscarriages. Hence, better molecular diagnosis is likely to become a cornerstone for investigation and treatment of unexplained losses of normal karyotype embryos. Nevertheless, the current data emphasize the role of activation of coagulation as a common final pathway to miscarriage mediated by placental vascular compromise.

Interventions to prevent pregnancy loss

Intravenous immunoglobulin

Although what we know or do not know with respect to how a treatment may work has no bearing on whether or not it does work (to paraphrase Sir Peter Medawar), it is generally believed that understanding underlying mechanisms in recurrent abortions allows a more focused, rational, and potentially more effective approach to treatment. Ultimately, the arbiter of such confidence must be the controlled clinical trial, but one must first have an intervention to evaluate and some means of selecting those most likely to benefit. On this basis, acetylsalicylic acid plus heparin or an anticoagulant therapy, and has been suggested for use in recurrent miscarriages associated with various serum autoantibodies that suggest autoimmune reactivity.

The results of clinical trials using IVIg for treatment of recurrent pregnancy loss have been conflicting, and at least four explanations for the controversy exist. The first involves differences in study design. In some studies (Coulam et al., 1995; Stephenson et al., 1998), IVIg was begun prior to conception, whereas in other studies (Christiansen et al., 1992; German RSA/IVIg Group, 1994), treatment was initiated after a positive pregnancy test, at 5–8 weeks gestation a time when the pregnancy had already been ‘selected’ for success to some degree. The second explanation involves differences in patient selection. The women included in the different studies were from different populations with potential differences in risk factors for pregnancy loss. Since 55% of losses may have been chromosomally abnormal (Coulam et al., 1996), only 45% would be expected to benefit from IVIg. To compare different trials, the proportion lost due to chromosome abnormalities must be known for both the treatment and control group. In the Coulam et al. (1994) study, the benefit of IVIg was confined to patients who had FHL/IUF-D-type miscarriages and no benefit was noted in those aborting blighted ova. Other trials have not diagnosed the type of loss so precisely, perhaps because the potential relevance of such information was not appreciated. The third explanation is statistical. Small studies risk false negative and false positive results—a point that will be discussed in more detail below. Although meta-analysis provides a method for combining small trials to obtain a data set of increased power, at present there are only 200 individual patients in that pooled set (Daya et al., 1999) which is insufficient to detect a small effect of IVIg given the heterogeneity among patients and trial designs. The fourth possible explanation is differences between IVIg preparations. IVIg is prepared from donor plasma pools, and the spectrum of antibody activities depends on the donors; anti-lipopolysaccharide activity can vary from batch to batch. Not all IVIg preparations used in the clinical trials have been shown to inhibit NK cells or cytokine production in vitro and/or in vivo in the recipients (Szederday et al., 1999; Nachbaur et al., 1997; Kwak et al., 1996; Ruiz et al., 1996).

Allogeneic leukocyte immunotherapy

The biological basis for allogeneic leukocyte immunotherapy to prevent miscarriage of normal karyotype has been outlined earlier in this review. The controversy generated by various clinical trials to test the efficacy of leukocyte immunotherapy in miscarriage patients has been considerable (Clark and Daya, 1991). All of the clinical trial data are empirical and try to answer the question, ‘Is the woman more likely to take home a baby if immunized with allogeneic leukocytes?’. Because it was not possible at that time to mount a multicentre randomized trial of sufficient size to have the statistical power to resolve the issue, this question was addressed by means of an international collaborative prospective observational study and meta-analysis by the Recurrent Miscarriage Immunotherapy Trials Group (RMITG) under sponsorship of the American Society for Reproductive Immunology (Coulam and Clark, 1991; RMITG, 1994). Individual data sheets from 1753 patients in published and unpublished randomized controlled or cohort-controlled trials were submitted for analysis to two independent statistical teams for meta-analysis (RMITG, 1994). This was done because it had been shown that randomized trial data published in the literature in other areas might not be reliable (Stewart and Palmer, 1993). Clarification of particular items was finalized at a Consensus Conference at which each of the data centres had to re-present their data (Proceedings of Consensus Conference, 1994). The RMITG meta-analysis study showed a 9–10% absolute improvement in take-home baby rate in all alloimmunized couples in randomized controlled trials. Double-blind randomized trials

Unexplained sporadic and recurrent miscarriage
showed a slightly greater efficacy compared to unblinded trials, contrary to what had been expected if there were a bias in unblinded studies (Clark et al., 1997) and cohort-controlled trials also did not show a greater benefit than randomized studies (RMITG, 1994). Logistic regression analysis showed no significant benefit of immunization for women with a previous live birth with their partner, or if the woman already had antipaternal HLA antibody (RMITG, 1994). Conversely, immunotherapy reduced the success rate in women who had autoimmunity [as reflected in a positive antinuclear antibody (ANA) or anti-cardiolipin (ACL) antibody]. Taking all couples in the study, one had to treat 10–11 women with antipaternal HLA antibody to achieve one additional live birth compared to using placebo; subset analysis examining paternal leukocyte immunization in primary recurrent aborters without any contraindication to treatment showed one only needed to immunize three to five patients with four or more prior losses (Daya et al., 1994). Although <2% of the abortuses had been karyotyped in the RMITG study, assuming a 55% expected rate of abnormality, chromosome errors were shown to be able to explain most of the failures after immunization (Clark et al., 1996; Clark and Coulam, 1996).

Recently, the result from a large multicentre National Institutes of Health-funded trial have been published (Ober et al. 1999). It had been decided in advance that analysis of the data should be done by intention to treat (Karrison and Ober, 1998), and this showed no statistically significant effects of treatment in primary plus secondary recurrent aborters or primary recurrent aborters examined separately; however, subgroup analysis of those who achieved pregnancy suggested a reduced success rate in immunized women ($P<0.02$). What is to be made of these data?

There have been two previous ‘negative’ trials of leukocyte immunotherapy (negative meaning the treated patients did worse than the controls). The first of these (Cauchi et al., 1991) attempted to reproduce the original Mowbray trial which used paired treatment and control couples in a sequential Armitage analysis (Mowbray et al., 1985; RMITG, 1994). The Cauchi trial used a much lower dose of immunizing cells, most women took a long time to achieve pregnancy (>3 months), and there was an imbalance despite randomization such that the treated patients had four prior miscarriages and the controls had had three prior losses (Clark and Daya, 1991; RMITG, 1994). The length of time to pregnancy was thought to be relevant since Mowbray had shown that only women making an antibody response had protection beyond 3 months (Mowbray, 1988); prior to that, antibody was unnecessary, a point against the relevance of ‘blocking’ antibodies. Boosting in antibody-negative women conceiving >3 months from primary treatment was found to restore protection (Mowbray and Underwood, 1991). Interestingly, boosting, or a primary immunization in pregnancy, had to be done by the 40th gestational day to be effective—an observation consistent with the idea that the immunologically modifiable lesion(s) causing miscarriages begin at the 6th week of gestation and not before. In the original Mowbray trial, boosting was not done, and when eventually all of the data were obtained for the RMITG analysis in 1993, it was found that the effect of immunization no longer achieved significance ($P<0.05$) (RMITG, 1994; Proceedings of Consensus Conference, 1994). The second negative trial was from Prof. Parazzini’s group, and when first presented it appeared that the women had had their abortions with different partners (i.e. the analyst could not confirm the losses were partner specific) (RMITG, 1994). These data were eventually published and the losses were stated to be partner specific (Illemi et al., 1994); however, the published data showed that pregnancies did not occur until >3 months from the time of a single immunization with $200 \times 10^6$ cells, and there were slightly more patients with four or more prior abortions in the immunotherapy group than in the control group (9/22 versus 6/22). The magnitude of the ‘negative’ effect of immunotherapy was enhanced by considering only those who became pregnant, rather than all who were given treatment (RMITG, 1994).

The meta-analysis of Collins and Roberts (RMITG, 1994) is updated in Figure 1, with extra data added from the original Mowbray trial, and the result from the Milan trial included, giving a total sample size of 501. The effect of adding in data from randomized controlled trials from two other centres where unfortunately it has not proven possible to obtain the primary patient data is also shown (RMITG, 1994; Clark et al., 1997), for a total sample size of 603. We now address the question of whether this sample size should be increased further by adding the data from Ober et al. (1999)?

There are a number of serious scientific problems with the Ober trial per se. As in all randomized controlled trials to date on leukocyte immunotherapy, the study size is small. Based on the eventual size (187 patients), given an expected 10% treatment effect shown by RMITG, there was a >60% chance that the Ober trial would give a non-significant result and a 10–11% chance of a ‘negative’ result where the immunized patients would appear less successful than the controls (Clark, 1999b). In the original RMITG study, from the 449 patients in the analysis, the risk of a Type II error was only 8.5%. It is standard design procedure to aim for a sample size that will detect the expected difference with a $P_a$ of 0.05 (5% risk of detecting an effect by chance) and power of 0.8 (1 – $P_B$, where $P_B$=0.2, a 20% risk of not detecting an effect which is real). Table I illustrates the importance of sample size. It can be seen that with a sample size of ~450 and slightly lower probabilities of success in the treatment (0.617) and control (0.517) groups, the RMITG result exceeded the minimum acceptable power value of 80%. Lack of analysis by centre (centre effect) further reduced the power of the Ober trial (Vamvakas, 1999).

As studies with small numbers of patients provide a less precise estimate of the actual result in the general population, they are more likely to give false positive and false negative results, and their $P_a$ values, whether <0.05 or >0.05, are insufficient to achieve credibility. The reason for a false negative is easily understood in terms of $P_B$, but the issue of false positivity is not as readily appreciated. Most consider the $P_a$ value to estimate the risk of a false positive, and 0.05 means a 1/20 risk. However, we search for positive results, a positive trial is ‘news’, and gets published. In fact, the one positive result could occur at a much higher probability by chance alone due to the phenomenon of multiple testing or multiple trials. As the number of trials is not usually known, we overvalue $P_a$. A similar problem underlies subset analysis, and the $P_a$ = 0.02 reported by Ober suggesting that immunotherapy is deleterious in one subset of patients is unreliable. Reproducing a result is important, but if a difference in treated patients is real at $P_a$ 0.05 (one-tailed), there is only a 50%
chance that the repeat trial will be significant. That is important to
know because we so often decide that a treatment does not work
when a repeat trial is negative. A two-tailed \( P \) value improves the
chance of a repeat trial being significant at the one-tailed level.
Indeed, the better the \( P \) value, the higher the probability of a
significant repetition of the result; small \( P \) values correlate with
increased power. The best way to address the repeat trial problem
is by meta-analysis of primary patient data from both published
and unpublished trials, as done by the RMITG (1994).

Although adding the data from the Ober trial to the existing
RMITG data set would appear highly desirable as a way of
addressing the problem, there are some major problems to such an
approach. Ober et al. (1999) did not indicate exclusion of partner-
non-specific recurrent aborters, those with ANA were not
identified and excluded, and although ACL positivity was an
exclusion criterion, the definition of ‘significant’ positivity
appears less sensitive than that used in the logistic regression
analysis done by the RMITG (1994) which showed a marked
reduction in pregnancy success rate if either ACL or ANA were
detectable. In the study by Ober et al., due to poor accrual, the
entry criteria were relaxed such that recruitment increased, and
many of these patients may have had subclinical antiimmunity
detectable by more extensive testing that included ANA.

Independent analysis of all the available laboratory test data by
two independent teams (RMITG, 1994) would provide the most
reliable method for evaluating the potential significance of
abnormal test results. It is known that recurrent miscarriage can
be a harbinger of subsequent development of autoimmune
disease, e.g. scleroderma (Silman and Black, 1988), but no data
are provided about possible early symptoms of autoimmune
problems in the Ober trial patients. There is also a curious result
in a post-hoc analysis (Ober et al., 1999) which was restricted to
those patients who achieved pregnancy. In the treatment group
there were seven patients with abnormal karyotype abortuses,
which, if removed from the analysis on the basis that such
problems were not immunologically modifiable, eliminated the
\( P < 0.02 \) that suggested reduced success rates following immu-
notherapy. It is possible that these occurred solely because,
as stated by Ober at the September 1999 ASRM symposium on
this topic, 17 of the 21 successful karyotypes were done in the
immunized patients; karyotyping was only achieved in four
patients in the control group (\( P < 0.025 \) by \( \chi^2 \)). The overall
success of karyotyping was only 36%.

Ober et al. claimed to have used the Mowbray immunotherapy
method in their trial. However, the immunizing cell dose of
\( 250 \times 10^6 \) cells was less than in the original Mowbray study that
used all the cells from 400–500 ml blood (Mowbray et al., 1985;
RMITG, 1994). Further, there was no analysis of outcome with
respect to time from immunization to conception in antibody-
negative (74% of sample) and positive women, and no boosting
until 6 months (Ober et al., 1999). Even if randomization had
generated two comparable groups of autoimmune-negative
partner-specific recurrent aborters, a major problem was intro-
duced into the Ober trial by the use of purified mononuclear cells
stored overnight before use. As shown in Table II, in the CBA/
JXDBA/2 system, even storage in serum-containing medium at

---

**Figure 1.** Meta-analysis of randomized controlled trials of allogeneic leukocyte immunotherapy in unexplained recurrent miscarriage patients. PUD = published and updated; UP = unpublished; UP/P = unpublished at time of 1993 conference, and subsequently published. Results for each study are given as mean (o) and 95% confidence interval. Pooled mean and 95% confidence interval (\( \sqrt{ } \)) is shown. Fixed effects Mantel–Haenszel model used (as in RMITG, 1994); a similar result was obtained using Peto, and DerSimonian and Laird (random effects) methods (Clark et al., 1997).
D.A. Clark et al.

Table I. Clinical perceptual problems arising from small clinical trials to detect a 10% real effect

<table>
<thead>
<tr>
<th>Trial</th>
<th>Success with treatment (%)</th>
<th>Success with control (%)</th>
<th>Difference (%)</th>
<th>$P_a$</th>
<th>Perception</th>
<th>$1-P_b$</th>
<th>Truth (power)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>33/50 (66%)</td>
<td>28/50 (56%)</td>
<td>10</td>
<td>&gt;0.05</td>
<td>Does not work</td>
<td>0.27</td>
<td>False negative, risk [equivalent]7/10</td>
</tr>
<tr>
<td>B</td>
<td>66/100 (66%)</td>
<td>56/100 (56%)</td>
<td>10</td>
<td>&gt;0.05</td>
<td>Does not work</td>
<td>0.43</td>
<td>Still false negative, risk &gt;2</td>
</tr>
<tr>
<td>C</td>
<td>132/200 (66%)</td>
<td>112/200 (56%)</td>
<td>10</td>
<td>&lt;0.025</td>
<td>Works!</td>
<td>0.66</td>
<td>Risk of false negative [equivalent]1/3</td>
</tr>
<tr>
<td>D</td>
<td>198/300 (66%)</td>
<td>168/300 (56%)</td>
<td>10</td>
<td>&lt;0.01</td>
<td>Works!</td>
<td>0.97</td>
<td>Risk of false negative 1/36</td>
</tr>
</tbody>
</table>

$^a$One tail $\chi^2$ with Yates’ correction. $P_a$ gives probability that a difference as large or larger than that observed could be obtained by chance alone. By convention, one accepts a 1/20 risk of making a fool of oneself. However, the risk that one’s colleagues will do a trial that disagrees with your result is much higher. Even with a trial of 200 per group, in a series of repeat trials 26% (1/4) will fail to give a $P_a<$0.05 notwithstanding a true/real 10% benefit. To minimize one’s risk, it is necessary to calculate power ($1-P_b$).

$^b$For a 34% expected reduction in abortion rate and number of participants per group, power ($1-P_b$)>0.99. Meta-analysis of six independent experiments (trials) has confirmed loss of protective activity with overnight storage at 4°C.

$^c$A power of 0.8 ($P_b=0.2$, 20% risk) is the minimum level currently accepted for trial design. A false negative risk also estimates the likelihood that a result with $P_a<0.05$ will not be achieved in a repeat trial containing the same number of subjects in each group. It is risky to decide treatment policies based on a single negative trial where the power is <0.8.

4°C eliminates efficacy. Storage at room temperature would have introduced the additional problem of production Th1 cytokines by the stored leukocytes, as shown in humans (Aye et al., 1995; Heddle et al., 1999). These Th1 cytokines would be expected to prevent a Th1→Th2 shift in vivo, and might even promote abortogenic Th1 responses in the recipient (Clark et al., 1999b). There have been two groups who claim success using stored blood leukocytes for immunotherapy of recurrent miscarriages. One group (Taylor et al., 1988) used $>10^9$ donor leukocytes, and another (Unander, 1988) used multiple units of erythrocyte-compatible whole blood transfusions, again $>10^9$ cells. Neither of these studies included a control group. The leukocytes in stored blood are thought to produce a transfusion-related immunomodulation that may increase tumour growth and susceptibility to post-operative infection (Vamvakas, 1999; Vamvakas et al., 1999). There are $>2500$ patients who have been studied in randomized trials of leukodepletion. A dose–response effect has been noted (Jensen et al., 1996; Houbiers et al., 1997); it is estimated for the dose-outcome effect in the trials that immunomodulation by stored cells requires $>2$ units of cells. The fact that 26% of the women immunized by Ober et al. developed antipaternal antibodies does not necessarily indicate that their immunization induced the cellular immune changes that are associated with protection against abortion; this protection has been shown to be independent of detectable antibodies. Thus, it would seem that Ober et al., while administering paternal antigen, were not administering cells that could achieve a protective effect. We have polled contributors to the RMITG study, and all respondents have indicated that they used freshly isolated leukocytes. Exactly what is altered by storing the cells is under investigation. Our preliminary data suggest that the transfusion-mediated immunomodulatory effect may depend on a subset of blood leukocytes expressing the tolerance signal OX-2 (CD200), which is depleted by overnight storage even at 4°C (Clark et al., 2001b).

Randomized clinical trials are time-consuming and expensive. It is not uncommon to discover a fatal flaw after the trial has been completed (Sackett and Hoey, 2000). By the time a trial of a particular treatment has been completed, that treatment may be outmoded; however, when putting a particular method that is well described to the test (Mowbray et al., 1985), it is not acceptable to change the recipe (immunization dose (Cauchi et al., 1991); immunization dose and leukocyte pretreatment (Ober et al., 1999)). Further, with uncommon problems, power achieved by adding patients that do not belong (because they do not have the problem) is no power at all. As the number of subjects in an analysis increases, the effect of heterogeneity will eventually approach zero, if randomization is properly done, but the magnitude of the effect shown in Figure 1 still requires a number needed to treat for one live birth as about 10 patients. This is as good if not better than many ‘medical’ therapies in use in other conditions (Clark and Coulam, 1996), but it is important to realize that if one could eliminate problems that cannot be corrected by immunotherapy, such as recurrent chromosome errors in the embryo (Clark et al., 1996), a 90% success rate would be expected such that for every 11 women treated, 10 would have a livebirth.

Might paternal leukocyte immunotherapy do harm? There have always been concerns about virus transmission, future difficulties matching for a graft or blood transfusion, and possible effects on the baby of antibodies to paternal antigens on erythrocytes, platelets, and granulocytes, and a cogent argument can be made for adequate purification of the mononuclear leukocytes and avoidance of use ofuffy coat cells (Clark and Coulam, 1996). Some patients do not achieve a pregnancy after leukocyte immunotherapy. Might allogeneic leukocyte immunotherapy shift endometrial cytokines towards reduced production of leukaemia inhibitory factor (LIF) and increased production of TNF-α so as to produce infertility? Reduced LIF and reduced Th2 cytokine
production plus increased Th1 cytokines are linked to infertility and spontaneous abortions (Piccinni et al., 1998) in patients who have not had immunotherapy, and if the underlying LIF deficiency were becoming progressively worse as part of the natural evolution of the problem, it would a priori seem unlikely to be a result of alloimmunization. There are some data to test the hypothesis. In the RMITG study (1994), 21.5% (219/279) of immunized patients in randomized controlled trials did not conceive, but neither did 25% (186/248) in the control groups. To show a significant (P<0.05) reduction in fertility would have required a non-conception rate in immunized patients of 30.1% for an average sample size of 264. The power is ~0.52 (1 – Pβ), suggesting a 48% risk of missing a significant effect with this sample size. From Table I, it would seem risky to accept the null hypothesis that there is no effect on fertility. Further, the data reflect what happened to a group of patients, and one cannot exclude that immunotherapy might render some patients more fertile and others less.

There are two possible solutions to such worries. One could merely throw up one’s hands and let ‘God’s will’ prevail; an act of omission is less likely to invite a lawsuit than an act of commission, and so far, even in the USA, it has not proven possible to obtain damages from the Almighty, although it has been possible to get large settlements for such problems as leaking silicone breast implants where the scientific evidence indicates such leaks have no effect at all. If the Hippocratic admonition of primum non nocere were to prevail, notwithstanding evidence that a treatment overall did more good than harm, there would be no drugs prescribed, no curative or palliative cancer chemotherapy, no coronary artery surgery, no IVF done. Alternatively, one could begin to analyse properly the cell and cytokine properties of the endometrium of recurrent aborters, their decidua when pregnancy fails, and the effect of immunization of endometrial cells and cytokines by pre- and post-treatment assays. This is akin to the bacteriologists’ approach to someone with fever, cough, and an infiltrate on chest X-ray; instead of giving everyone penicillin (to which most will not respond and some only with dangerous allergic reactions), the approach is to identify who is likely to respond by identifying the pathogenic agent, its susceptibility to the treatment, documenting that that treatment produces bacteriological cure, and watching carefully for side-effects. Such an approach means that miscarriage patients will continue to need to be studied in centres with the specialized technology needed for cellular and molecular analysis of their uterine lining, as well as being able to detect over-expression of fgl2 procoagulant in early pregnancy. One may eventually be able to extend these types of studies to prevent the first miscarriage. A suitable approach allowing treatment that can do good and possibly harm is to ensure that each patient and her spouse have provided written informed consent. Meanwhile, neither clinical trials nor clinical immunotherapy with allogeneic leukocytes should be undertaken by non-specialists.

New options for the new millennium?

Both IVIg and leukocyte immunotherapy are expensive and time-consuming approaches. The benefits of psychological support and/or formal psychotherapy have been suggested as an alternative, but not a single randomized trial has been done and the cost can be greater than immunotherapy (Clark and Daya, 1991; Clark, 1999a). Current emphasis on the role of clotting and inflammation in causing miscarriages has raised hopes that standard pharmaceutical preparations such as heparin, acetylsalicylic acid, and prednisone, purified cytokines such as IL-10, or defined bacterial antigens such as killed streptococcus could be used (Katano et al., 1997; Blumenfeld and Brenner, 1999; Clark, 1999a). One randomized trial (Laskin et al., 1997) reported that prednisone plus aspirin improved the successful pregnancy rate from 55 to 65% in patients who may or may not have had significant autoantibodies; this 10% difference is strikingly similar to that reported by RMITG (1994) (and Figure 1), but Laskin et al. concluded that there was no effect, a classical Type II error when sample size is too small. In a larger series of 678 patients (Reznikoff-Etievant et al., 1999), of whom 230 had detectable autoantibodies, there was a 90% success rate in 161 autoantibody-negative patients treated with aspirin plus short-term prednisone, significantly better than in 63 patients given aspirin alone; in 53 patients with autoantibodies who achieved pregnancy and were given aspirin plus prednisone, 85% successfully gave birth. With short-term (2 months) treatment, the side-effects seen by Laskin et al. (1997) did not occur. The 90% success rate reported by Reznikoff-Etievant could only occur if the population had a low incidence of chromosome abnormalities or treatment had some effect on this problem. Unfortunately, the patients who became pregnant in the 678 patient sample were not randomized! However, the fact that such a large sample can be generated strongly suggests one could conduct trials of sufficient size to determine if alternatives to IVIg and allogeneic leukocyte immunotherapy are efficacious and safe.

Conclusions

There have been significant advances in our understanding of the pathogenesis of ‘unexplained’ recurrent pregnancy loss. Similar mechanisms probably occur in women having a single unexplained miscarriage. Occult losses probably have a different mechanism than classical miscarriages. Allogeneic leukocyte immunotherapy using freshly prepared blood mononuclear leukocytes may be beneficial for a subset of women having recurrent clinical miscarriages of normal karyotype embryos beginning after 6 weeks gestation who have no evidence of autoimmunity. Patients who become pregnant >3 months from the time of primary immunization should be considered for boosting beginning after 6 weeks gestation. Randomized trials of alternative treatment strategies are needed. Several small trials wherein the primary patient data are audited and pooled in a meta-analysis provide a more reliable indication of efficacy than a single large trial, whether multicentre or not.

References


Unexplained sporadic and recurrent miscarriage


Received on September 15,2000; accepted on June 20, 2001