NEW DEBATE

Unexplained infertility: Does it really exist?

N.Gleicher1,2,3,5 and D.Barad1,2,4

1Center for Human Reproduction, New York, NY, 2Foundation for Reproductive Medicine, Chicago, IL, 3Department of Obstetrics and Gynecology, Yale University School of Medicine, New Haven, CT and 4Department of Epidemiology and Social Medicine and Department of Obstetrics and Gynecology & Women’s Health, Albert Einstein College of Medicine, Bronx, NY, USA

To whom correspondence should be addressed at: Center for Human Reproduction, 21 East 69th Street, New York, NY 10021, USA.
E-mail: ngleicher@thechr.com

Recent medical literature has quite extensively addressed the use of various terminologies within the field of reproductive medicine. This discussion has, however, so far overlooked the fact that one of the most frequently made diagnosis, so-called unexplained infertility (UI), not only didactically but, even more importantly, clinically, appears unsustainable as an independent diagnosis. The arguments in support of such a contention are manifold. The diagnosis of UI is highly subjective. It is dependent on which diagnostic tests have been performed (or have been omitted) and at what level of quality. Paradoxically, a diagnosis of UI will, therefore, be more often reached if the diagnostic workup is incomplete or of poor quality. Supported by evidence from the literature, the argument is made that the conditions, most frequently misdiagnosed as UI, are endometriosis, tubal infertility (especially distal and peritubal disease), premature ovarian ageing and immunological infertility. Because of the obvious unreliability of a diagnosis of UI and the widely reported unevenness in diagnostic criteria, we recommend the abandonment of UI as a formal infertility diagnosis. Better efforts to reach infertility diagnoses more accurately should improve the diagnostic accuracy of hitherto frequently missed diagnoses, which often falsely have led to a diagnosis of UI.

Key words: endometriosis/immunological infertility/ovarian ageing/tubal infertility/unexplained infertility

‘The surest way to convey misinformation is often to tell the strict truth.’

Mark Twain

Introduction

In attempts to improve communications between reproductive practitioners, the literature on both sides of the Atlantic, in recent years, has abounded in publications that have attempted to improve terminologies. Amongst these attempts, surprisingly little attention has, however, been given to a very frequent diagnosis that, already in name, reflects most of its shortcomings: unexplained infertility (UI) may, indeed, represent the single most frequent female infertility ‘diagnosis’, with a reported prevalence of approximately 25–30% of all infertility (Evers, 2002; Smith et al., 2003). Authoritative European sources have previously noted that ‘the diagnosis of UI is one of exclusion (and) it is therefore important to seek agreement on which diagnostic tests are required to be done before concluding that a couple have UI’ (Crosignani et al., 1993). The same authoritative body, however, also has stated that ‘such agreement is not readily obtained’ (Crosignani et al., 1993).

There are many reasons why reaching specific infertility diagnoses can be difficult: one can never be sure that, even a well-documented diagnosis, represents the only cause of a couple’s infertility. Other contributing causes may be unknown to the treating physician or, when known, be part of so-called multifactorial infertility. The presumptive diagnosis of UI is reached when whatever diagnostic workup has been chosen reveals no obvious cause for a couple’s infertility. The didactic term UI, therefore, does not, as diagnostic terminology usually does, describe a specific clinical condition, characterized by specific diagnostic findings, but is used to define a negative, the absence of specific diagnostic findings. In other words, the diagnostic terminology of UI is not used to describe the presence of a medical condition, but a void—a negative. Negatives are, however, practically impossible to prove. To quote Carl Sagan: ‘Absence of evidence is not evidence of absence’.

Words have very specific meanings: if UI were simply to be used as a universally accepted terminology to describe such a diagnostic void (in other words, our inability to reach a diagnosis, whether uni- or multifactorial), the term would be probably acceptable. Within the currently existing context, UI, however, stands as a specific infertility diagnosis listed in medical textbooks on a par with other standard clinical infertility diagnoses.

Let us, for a moment, assume that UI, indeed, were to exist. Such a diagnosis should then only be reached if all appropriate diagnostic tests were performed and have failed to detect one or
more presumed causes for a couple’s infertility. What represents ‘all’ appropriate diagnostic tests has, however, remained undefined, as the literature is quite unanimous in suggesting that there is no agreement as to what constitutes a complete infertility workup [Crosignani et al., 1993; ESHRE Capri Workshop Group, 1996, 2000; American Society for Reproductive Medicine (ASRM), 2000]. What constitutes an appropriate infertility workup may, in addition, also vary based on clinical circumstances. For example, age of females will, of course, greatly affect the decision-making process (ASRM, 2002). In addition, it seems obvious that the quality of performance of diagnostic procedures can vary and, therefore, with it, their reliability. Finally, the diagnostic relevance of tests may be interpreted differently by different clinicians. For example, the evaluation of antiphospholipid antibodies in infertile patients has been strongly discouraged by some (ASRM, 1999). Yet, others disagree, based on reports which have demonstrated that subclinical autoimmune disease, already in the preclinical stages of classical autoimmune diseases, has been associated with decreased fecundity (Nelson et al., 1993; Sicman and Black, 1998; Gleicher, 1999). Abnormal (auto)immune function is also found at increased prevalence in infertile women (Gleicher et al., 1989; Geva et al., 1997).

Another example for the high degree of observer-to-observer variation in the interpretation of diagnostic findings has been well documented, in women with endometriosis. Endometriosis is frequently misdiagnosed as UI (Cook and Rock, 1995), because a diagnosis of endometriosis, even with laparoscopy, can be easily missed (Olive and Schwartz, 1993). Similarly, the diagnostic accuracy of hysterosalpingography (HSG), a mainstay of infertility diagnosis, has been questioned (Gleicher et al., 1992).

To account for such disagreements about the validity of diagnostic tests, European colleagues have suggested dividing such tests into three categories: those with established correlation with the occurrence of pregnancy, those not consistently showing correlation with the occurrence of pregnancy and those apparently not correlating with the occurrence of pregnancy at all. They, furthermore, concluded that, except for laboratory assessments for ovulation, assessment of tubal patency and semen analysis, ‘... other additional investigations contribute relatively little to effective diagnosis of unexplained infertility and so are not mandatory’ (Crosignani et al., 1993).

Such a dismissive attitude in respect to the value of many diagnostic procedures is quite prevalent. For example, in an otherwise outstanding review, Evers (2002) rejects the use of immunological testing. Although he may be correct in his opinion, he based most of his argument on one single study that could not statistically differentiate between 32 fertile and infertile couples, a quite obviously inadequate sample size (Guzick et al., 1994a). This example is noted, not to take sides as to whether immune parameters should be tested for but to demonstrate the difficulties involved in reaching consensus on what, indeed, does represent adequate diagnostic testing.

To maintain the current status quo and to accept the widely held concept that a diagnosis of UI is appropriate, as long as ovulation is confirmed, tubal patency has been proven and the semen analysis is normal, therefore, remains as an option. Acceptance of such an option would, however, mean acceptance of the fact that our most frequently made diagnosis in infertility practice is a non-diagnosis.

Such an approach appears, however, no longer acceptable. As the last two decades have demonstrated (and is further discussed below), even apparently normally ovulating women may produce poor quality oocytes and embryos, patent fallopian tubes do not necessarily function normally and various subleties in superficially normal semen analyses can affect fertilization and, with it, fertility potentials. It would, therefore, appear that, considering our current knowledge level, the diagnosis of UI should no longer be based on the use of a catch-all diagnostic category, which really only represents a statement of diagnostic ignorance and, in such a sense, is just repetitive of diagnostic platitudes of the past.

The potential difficulties in reaching one, or more, infertility diagnoses do not only lie with ordering the correct tests and performing them appropriately. Equally important is, of course, the correct interpretation of diagnostic findings. In the remaining parts of this discussion, we will address, what we believe to represent, some of the principle causes of misdiagnoses in infertility and, why, in our opinion, such misdiagnoses can greatly contribute to the (mistaken) non-diagnosis of UI. We fully recognize that some of our opinions currently still lack the support of well-controlled studies, which could be considered as the highest levels of evidence. However, the generation of new evidence is a step-wise approach. By presenting what we consider the best available evidence, even if not always perfect, it may contribute to the conduct of studies, which, in the end, will give us the highest levels of evidence.

The principle causes of misdiagnoses

As previously noted, a specific cause of infertility is never certain. Even in situations of certain sterility (in contrast to infertility), a couple’s ultimate failure to conceive may have additional causes. For example, the couple with quite obvious infertility, because of azoospermia, may still not conceive with inseminations with normal sperm, if the female’s fallopian tubes are blocked. This is the reason why initial diagnostic workups always encompass both partners, even if one partner presents with an obvious infertility problem. The complexities involved, in reaching correct diagnoses, have, therefore, to be acknowledged, and the literature provides considerable evidence that final diagnoses are often inaccurate. The four below-discussed medical conditions appear to contribute most often to the non-diagnosis of UI.

Endometriosis

The prevalence of endometriosis in infertile females has remained controversial and has been reported to be as low as 5–10% (Smith et al., 2003) and as high as 30–50% (ASRM, 2004). Even in the hands of experienced laparoscopists, an accurate diagnosis can be difficult because the disease is often microscopic in nature and presents visually with a variety of atypical lesions (Olive and Schwartz, 1993; Cook and Rock, 1995). Many investigators have pointed towards similar patient profiles in women with mild cases of endometriosis and UI. Suggestions have, therefore, been made that endometriosis
may be underdiagnosed and that UI may, in many cases, represent a non-visible and/or only microscopic precursor stage of endometriosis (Guidice and Kao, 2004).

Some authorities have argued that the mere presence of mild endometriosis, in the absence of visible secondary organic disease, for example tubal involvement, is not associated with infertility (Koniwckx, 1994; Moen, 1995). Consequently, they consider a diagnosis of UI in such cases as appropriate. We disagree with this argument: there is now overwhelming, recently well-summarized, evidence in the literature that endometriosis affects IVF outcomes in practically all aspects of the process (Guidice and Kao, 2004). If endometriosis does so in vitro, it can be expected to exert the same effects in vivo, as well. Further evidence for a significant impact of, even mild, endometriosis comes from earlier data on gamete intra-fallopian tubal transfers (GIFTs) of embryos. When GIFT still represented standard of care, a number of studies demonstrated that, even in mild cases of endometriosis, pregnancy outcomes were impacted by subtle tubal abnormalities (Fakh and Marshall, 1994; Guzick et al., 1994a) This observation is, of course, reflective of microscopic endometriotic disease in fallopian tubes (and other organs), which can never be fully ruled out, even by laparoscopy (Oliver and Schwartz, 1993; Cook and Rock, 1995).

The logical conclusion from these findings, therefore, has to be that mild endometriosis can affect normal tubal function, as well as other reproductive processes (Guidice and Kao, 2004). If one accepts that this diagnosis is clinically frequently overlooked, and if no other obvious cause of infertility is apparent, the patient will end up with the incorrect non-diagnosis of UI.

**Tubal disease**

The inaccuracy of HSG in defining normal tubal physiology and anatomy has been well established. Consensus exists that HSG is less accurate in detecting and evaluating tubal disease than laparoscopy (Opsahl et al., 1993; Mol et al., 1999; Tanahatoe et al., 2003), is especially poorly suited to assess distal tubal disease and peritubal disease (Swant et al., 1995; Mol et al., 1996; Glattstein et al., 1997), suffers from considerable observer variability (Glattstein et al., 1997) and can be improved upon by modifying standard HSG techniques (Gleicher et al., 1992; Karande et al., 1996; Woolcott et al., 1999).

As previously noted, despite these very obvious inadequacies, HSG is widely considered a corner stone of infertility diagnosis and still represents the principle first line tool to assess tubal status (Crosignani et al., 1993). Yet, as the consensus shows, routine HSG has limited value in assessing tubal patency and basically no value in determining tubal function. Tubal function can be impaired in the presence of tubal patency, as has been well documented (Karande et al., 1996; Papaioannou et al., 2003). Indeed, we demonstrated that, in an infertile population, routine HSG misses at least one anatomical or physiological tubal abnormality in 84% of the cases (Karande et al., 1995a) and that the detection of the physiological function abnormality of an elevated tubal perfusion pressure is, in 85% of cases, associated with the laparoscopically confirmed presence of endometriosis (Karande et al., 1995b).

These studies are well associated with the previously noted high prevalence of tubal abnormalities in endometriosis patients undergoing assisted reproductive technologies (ARTs) (Fakih and Marshall, 1994; Guzick et al., 1994a; Guidice and Kao, 2004). The obvious conclusion from these data is, of course, that at least some women with a diagnosis of UI actually are infertile because of undiagnosed tubal disease, often secondary to undiagnosed endometriosis.

**Prematurely ageing ovaries**

A number of investigators in recent years concluded that some women follow a premature ageing curve (Beckens et al., 2002; Nikolaoou and Templeton, 2003; Gleicher, 2005). This concept of prematurely ageing ovaries (PAO) is based on a number of observations: first, the process of ovarian ageing appears to be statistically correlated with the total number of remaining follicles within the ovary (te Velde et al., 1998; Faddy, 2000). The original supply of follicles continues to decline from birth. te Velde et al. (1998) noted that a state of subfertility is reached at approximate age 30–31 years, when the remaining follicles have become a fraction of their original number. By age 37–38 years, a critical point is reached, with approximately 25 000 follicles remaining in the ovaries. At that point, follicular depletion accelerates towards menopause, which is reached when the follicular count reaches approximately 1000, at an average age 51 years (Guzick et al., 1994a; te Velde et al., 1998; Faddy, 2000; Nikolaoou and Templeton, 2003). A second, very important observation suggests that the time period between accelerated decline in fertility (i.e. age 37–38 years and 25 000 follicles) and menopause (i.e. age 51 years and 1000 follicles) is fixed at approximately 13.5 years (Faddy, 2000; Nikolaoou and Templeton, 2003). As Figure 1 demonstrates.

![Figure 1](https://academic.oup.com/humrep/article-abstract/21/8/1951/2938641)  
Figure 1. The concept of the prematurely ageing ovary (PAO). Modified from Gleicher (2005) with permission.
these observations suggest that a woman, who reaches her acceleration point of fertility decline (of 25,000 follicles) early, will, therefore, most likely, also reach menopause early. In other words, such a woman ages her ovaries along an ovarian ageing curve which is shifted towards the left (Figure 1) but remains predictable as to when she will reach the various key points in fertility decline, including menopause. The more the curve is shifted towards the left, the earlier will the accelerated decline in fertility begin and the earlier will such a patient experience menopause (Gleicher, 2005).

The literature suggests that approximately 10% of women experience early menopause before age 45 years and 1% before age 40 years (van Noord et al., 1997). These women can be assumed to have followed a premature ovarian ageing curve based on which they hit points of subfertility (age 31–32 years) and accelerated decline (age 37–38 years) early. In other words, these women would present with evidence of diminished fertility when nobody would expect such a decline, based on their age.

For lack of symptoms, such women will frequently and erroneously be tagged with a diagnosis of UI (Nikolaou and Templeton, 2003, 2004). One, of course, would expect to see them disproportionately concentrated in infertility practices.

Women with PAO can be correctly diagnosed if alleged UI patients are carefully investigated for signs of PAO. The most reliable of such signs are age-atypical resistance to ovarian stimulation with gonadotrophins, (Beckens et al., 2002; de Boer et al., 2002, 2003; Nikolaou et al., 2002; Nikolaou and Templeton, 2003) and prematurely elevated baseline FSH levels, abnormal ovarian function tests, low antral follicle counts and a family history of early menopause (Nikolaou and Templeton, 2003, 2004; Gleicher, 2005).

At times, only an IVF cycle will reveal the full extent of PAO and, IVF has, therefore, been proposed not only as an ultimate therapeutic modality for infertile patients but also as the ultimate diagnostic test (Nikolaou and Templeton, 2003; Gleicher, 2005).

Subclinical autoimmune disease

We have in the Introduction to this article alluded to the controversies that still surround the possible association between abnormal autoimmune function and female infertility. Prominent investigators, as well as authoritative professional organizations, have published statements that deny the value of diagnostic efforts to detect subclinical autoimmune abnormalities in infertile patients (ASRM, 1999; Hill, 2000; Hornstein, 2000; Scott, 2000). Yet, these opinions are based only on the observation that proposed treatments failed to demonstrate outcome benefits. This, of course, does not, by implication, mean that all laboratory testing should be considered useless. The absence of good treatment options does not necessarily mean that abnormal autoimmune function may not be aetiologically associated with infertility. Indeed, as previously noted, there is considerable evidence in the literature to suggest such an association. Many investigators have reported a clustering of subclinical autoimmune abnormalities in infertile populations (Geva et al., 1997; Gleicher, 1999). The most convincing evidence comes, however, from the widely reported observation that women with classical autoimmune diseases, even before they reach diagnosis (i.e. become clinically symptomatic)—in other words, at pre-clinical stages of their impending autoimmune diseases—already demonstrate decreased fecundity (Nelson et al., 1993; Sicman and Black, 1998).

One, of course, can question the purpose of a diagnosis if it cannot be followed up with treatment. The hostility, expressed by many towards immunological investigations in infertile patients, can, therefore, be understood. Yet, on an analytical level, one, nevertheless, has to assume that at least some cases of UI may be a reflection of decreased fecundity because of abnormal immune function. The fact that we do not understand why abnormal autoimmune function reduces fertility and that we lack successful treatment options should not allow us to ignore the statistically well-established association between subclinical autoimmune diseases and decreased fecundity. To define such patients as unexplained would, therefore, appear intellectually dishonest.

Conclusions

We have presented evidence that the non-diagnosis of UI is, at best, highly subjective and inaccurate and, at worst, non-existent. Indeed, were it to be preserved, one would, at a minimum, have to accept the notion that a diagnosis of UI always has to be seen as provisional and subject to revisions. Any serious attempt to improve the didactic terminology in the field of reproductive medicine, would, however, suggest that there is no longer a place for such a diagnostic entity. A better effort should be undertaken to develop reliable tools to diagnose, hitherto often undiagnosed, conditions of endometriosis, tubal disease, premature ovarian ageing and immunological infertility, which are often misdiagnosed for UI. Diagnostic terminology should be based on evidentiary diagnostic findings and not on their absence.

Even the best, and most accurate, diagnostic approach will leave some patients without a specific diagnosis. We would argue that such patients should be considered 'undiagnosed', rather than have a formal diagnosis of UI attached to them. After all, they do not have a diagnosis!

References


