Enough! Stop the arguments and get on with the science of natural killer cell testing

Gavin Sacks\textsuperscript{1,2,3,*}

\textsuperscript{1}Institute of Medical Science, University of Sydney, Sydney, Australia \textsuperscript{2}St George Hospital and Royal Hospital for Women, Sydney, Australia

*Correspondence address. E-mail: gsacks@nchem.com.au

Submitted on March 25, 2015; accepted on April 7, 2015

ABSTRACT: Natural killer cell testing is currently practiced widely, and there are studies indicating potential benefit in terms of targeting women with repeated reproductive failure for immune therapy. This may be a better approach than empirical immune therapy without any investigation. More and better studies are needed before such an approach can be fully endorsed. There is still uncertainty over the precise pathophysiological basis for all immune investigation and therapy, but this should not be a barrier for clinical observation and empirical care. On the contrary, clinicians and researchers should work more closely together to provide the best care for our patients.

Key words: uterine natural killer cells / peripheral blood natural killer cells / IVF failure / recurrent miscarriage / reproductive immune therapy

Introduction

The remarkable division of opinion on natural killer (NK) cell testing described by Moffett and Shreeve (2015) is confusing for clinicians and often leaves patients poorly managed one way or another. So why is this one element of the overall maternal immune response to pregnancy cause for such controversy? The reasons for this are (i) NK cells are the main immune cells at the maternal–fetal interface, (ii) NK cells have the potential to orchestrate the overall immune response and, either directly or indirectly, influence trophoblast invasion, (iii) NK cell testing has a long history (20 years) of clinical use and is the most commonly available cellular immune test, (iv) clinical application is based on a simple hypothesis that does not account for our current level of understanding of NK function, (v) immune therapy for high NK cell levels has been promising in many studies, although some claims for benefit have almost certainly over-exaggerated the likely impact.

The debate: does NK cell testing have any clinical value?

Moffett and Shreeve (2015) approach NK testing from the basic science perspective, with a state-of-the-art summary of current knowledge of NK cell biology. But let us, for a moment, consider a different view. Imagine the perspectives of our patients with multiple miscarriages or IVF cycles, with relatively short and finite times to succeed, and each failure associated with significant physical, financial and emotional stress. It is a medical imperative for their clinicians to try whatever they can to minimize this morbidity. So the debate is actually very clear. Do we approach NK testing and treatment from a mechanistic pathophysiological framework, or from an empirical clinical one? Both have scientific validity. It is not necessarily the case that all insights come from the former. Clinical observations—even case reports—can sometimes be more insightful than any laboratory experiment or large randomized trial. If empirical clinical practice is suggestive of something, there is no point in simply claiming it is impossible based on current knowledge. It is wiser to try to test the claims, to reproduce the effects and to alter the current theoretical paradigm accordingly.

So I believe that this debate is actually a philosophical one rather than a scientific one. In brief, Moffett and Shreeve (2015) believe that not enough is known about NK cells in reproduction to use NK testing clinically, that clinical tests oversimplify a complex and more likely beneficial function in the establishment of early pregnancy, that women are being misled and mistreated by undergoing these tests and that they could be harmed by the associated treatments. Moffett and Shreeve (2015) conclude that it is surely no longer acceptable for licensed medical practitioners to continue to administer and profit from potentially unsafe and unproven treatments, based on belief and not scientific rationale. So let us consider these arguments in turn.

What is NK cell function in reproduction?

Given (i) the original description of NK cells as a first line of defence against invasion and (ii) that one of the main features of trophoblast cells is the absence of classic major histocompatibility complex proteins, it was naturally assumed that uterine NK (uNK) cells potentially pose a...
threat to the invading trophoblast (Sacks 1999; Chazoust, 2013; Clark 2014). Moffett and others have since demonstrated that the reality is indeed far more complex, with the majority of uNK cells being of a subtype with negligible cytotoxicity (Moffett and Colucci 2014), and 'activation' of those cells seems to confer better implantation rather than worse (Hiby et al., 2008; Xiong et al., 2013). These CD56 'super bright' uNK cells surround the conceptus (Moffett and Colucci 2014), perhaps even themselves preventing other more potentially harmful cells from getting near. But, as Moffett and Shreeve (2015) state, 'the actual contribution of these maternal immune cells to success or failure is still unknown'.

This view is based largely on research on CD56 super bright cells, which make up 90% of uNK cells. But what of the other 10% of CD56 'dim' cells? What if the proportion of uNK subtypes is altered? Can the control of implantation be lost by a more subtle alteration of the NK subtype ratio? The practical difficulty of examining such tissue means we currently do not know, although some studies using flow cytometry have demonstrated such an effect (Larchepelle et al., 1996; Fukui et al., 1999). Women with recurrent miscarriage do appear to have higher numbers of the more cytotoxic CD56dim subtype (and less of the CD56 super bright cells) even if overall numbers are unchanged.

It is now well known that uNK cell numbers vary enormously through the menstrual cycle (Russell et al., 2013), so there is significant potential for altered uNK subtype levels. Moffett and Shreeve (2015) describe two mechanisms by which uNK cell numbers can increase from in utero resident CD34+ stem cells (Vaccari et al., 2011), or from immature circulating progenitors (Male et al., 2010). But they conveniently do not mention the possibility of trafficking of mature peripheral blood NK (pNK) cells (Cattino et al., 2008; Kalkunte et al., 2009).

Moffett has previously argued that uNK cells are as different from pNK cells as black taxicabs in central London are to red minicabs driving around the city's orbital motorway (Moffett et al., 2004). This is true in a general sense only: 90% of pNK cells are CD56dim and 10% are CD56bright, whereas for uNK cells, it is the other way round. As is immediately obvious, there is overlap. Monthly trafficking of the more dangerous CD56dim cells could have a substantial impact on overall uNK function, and this hypothesis has simply never been adequately explored. Moreover, tests of pNK cells may potentially detect high levels of activated cells (CD56dim) which, due to their increased activity, may be more likely to be recruited into the endometrium each month. Thus, there is a theoretical link between blood and uterine NK cells, and there is indeed some evidence of an association between uNK and pNK cell numbers (Park et al., 2011; Sacks et al., 2014; Santillan et al., 2015).

Moffett and Shreeve (2015) describe how uNK cells could control the depth and pattern of trophoblast invasion, with observations that activation (and not inhibition) of uNK KIR receptors appears to promote effective implantation (Hiby et al., 2008; Alessandra et al., 2014; Moffett et al., 2015). But it is more complex still, with evidence for uNK control of vascular remodelling in early pregnancy (Robson et al., 2012). It has been reported that the absence of uNK cells leads to an exaggerated endovascular trophoblast effect on spiral arteries (Chakraborty et al., 2011), whereas impaired vascular remodelling in women with recurrent miscarriage has been associated with more irregular and dense cytoplasmic granules in uNK cells (Almassy et al., 2015). And, in further apparent contradiction, high levels of pNK cells have been associated with decreased uterine blood flow in early pregnancy (Yi et al., 2014). It is also likely that uNK cells are involved in other important events in the decidua. It has been shown recently that high levels of uNK cells are associated with in vivo corticosteroid deficiency (Kuroda et al., 2013). Perhaps the detection of high uNK cell numbers in some women has nothing whatsoever to do with 'immune rejection' but can signify a potential deficiency for therapeutic intervention.

Moffett and Shreeve's (2015) argument that we do not know enough can be applied to almost every area of medicine. It is wrong of them to deny the possibility that there is a rationale behind altered uNK status and function affecting implantation. It is obvious that the maternal immune response to pregnancy is complex and has a number of components (Sacks 2015). We are all agreed that we do not know the relative importance of each, or how they interact. But it is noteworthy that one of the authors has previously stated that 'the theory that patients benefit from suppressing pro-inflammatory Th1 cytokine activity ... remains plausible in our opinion' (Shreeve and Sacks 2012). So Moffett and Shreeve (2015) are not really against the concept of immune testing and treatment, but NK testing specifically. Let us examine what NK testing really is.

What are the tests?

NK testing is actually a lot more technically difficult than is often appreciated. It is easy to assume, as is inevitably done in recent attempts to create meta-analysis from current data (Jesudhin and Sinkara, 2014), that the tests are all the same. But they are not the same from clinic to clinic, there is no consensus on reference ranges and even studies reporting similar laboratory assays are based on different heterogeneous patient populations. So, rather than bland attempts to put them all in one convenient box and try to claim a greater truth from the perspective of 'evidence-based medicine', more attention needs to be paid to the tests themselves. Moffett and Shreeve (2015) highlight the dangers of sending blood samples overseas without any known control data. Anecdotally, the extremely high rate of 'abnormality' in those test reports is simply not believable or credible and is perhaps one reason for the strong opinions expressed by Moffett and Shreeve (2015).

Uterine biopsies for uNK testing are often assumed by clinicians to be the best assessment of NK activity and immune issues (Quenby et al., 1999; Drury et al., 2013; Santillan et al., 2015). This could not be further from the truth for two reasons. First, as so eloquently described by Moffett and Shreeve (2015), it is clear that simply measuring uNK cell numbers tells us nothing about subtypes and overall immune activity. Indeed, 90% of uNK cells are thought to be beneficial for establishment of pregnancy. And second, the technique of immunohistochemistry (staining cells for CD56 and counting down a microscope) cannot assess CD56 subtypes, and the counting is extremely subjective. This is particularly difficult with the glandular tissue examined, where discrete pockets of NK cells can give completely different results if included, or not, in the count (Russell et al., 2011; Manseh et al., 2012). In Sydney (Russell et al., 2011; Russell et al. 2013) and elsewhere (Drury et al., 2013; Lash et al., 2014; Santillan et al., 2015), efforts have been made to create a reasonably reliable counting system, but it is clear that this kind of work should not be undertaken lightly and without considerable work-up before attempting to use the data clinically.

Analysis of uNK cells is also hugely influenced by the day of the cycle in which the sample is taken. Again, most previous research studies have
underappreciated this. The increase in uNK numbers is so dramatic (e.g., from 5 to 40% of stromal cells) over just 10 days that it is essential for a laboratory to create reference ranges for each day. So far, as far as we know, we have been the only group to have achieved this (Russell et al. 2013). What this tells us is that the vast majority of previous research in which ‘high’ levels of NK cells were described for samples taken on a range of days (e.g., 4-7-9) is less informative than hitherto believed (e.g., Querby et al. 1999; Tuckerman et al. 2007; Tuckerman et al. 2010; Tang et al. 2013). I believe this whole area needs to be revisited with better control of cycle day of testing and the diagnosis of ‘high’ uNK levels.

The development of analysis of pNK cells in women with reproductive failure has continued in a virtual parallel universe to uNK testing, with the main link in name only. Indeed, Moffett and Shreeve (2015) in their article title mention ‘uterine’ NK cells specifically and incorporate pNK testing in their text as if it is simply an alternative means of assessing uNK cells. This is not the case. Yes, there are potential associations in that (i) pNK cells may be recruited to the uterus every month, (ii) there is a hypothesis that higher numbers of CD56dim cells from blood could alter the uNK subtype ratio and (iii) there are demonstrations of correlation between uNK and pNK cell numbers, made by us (Sacks et al. 2014) and others (Park et al. 2011; Santillan et al. 2015). But in truth, pNK testing does not need uNK testing for any validity at all. It has evolved in its own way, with its own rationale and observational studies. There are a number of different tests for pNK cells, and none bear any relationship with immunohistochemistry used for uNK testing.

Blood NK cell testing has been described using in vitro functional assays, and flow cytometric assays of overall cell counts, subtype ratios or expression of activation markers. Furthermore, over 20 years, a number of groups from around the world have reported higher levels of pNK cell activity in women with reproductive failure (Thurn et al. 2004; Thurn et al. 2005; Thurn et al. 2008; Chernyshov et al. 2010; King et al. 2010; Winger et al. 2011; Karani et al. 2012; Liang et al. 2012; Sacks et al. 2012; Dongkai et al. 2014; Ota et al. 2014).

Higher levels have been reported in women with recurrent miscarriage who subsequently miscarried again with a normal karyotype (Yamada et al. 2003). Higher levels of pNK cells in reproductive failure appear to be specifically in the CD56dim subtype rather than the CD56bright subtype (Mattoo et al. 2007). And this supports the hypothesis described above that, perhaps in such cases, recruitment of CD56dim cells into the uterus alters the immune environment. It is not yet proven that these levels are predictive of future performance or that targeted immune therapy intervention is beneficial. However, there are clinical studies supportive of these hypotheses.

Clinical studies: is there any evidence?

NK testing was first introduced to patients in 1995, with perhaps excessive claims associated with relatively poor quality observational studies (Beer et al. 1996). Alan Beer promoted the concept with a book (‘Is your body baby friendly?’) which continues to interest patients to this day. But rather than a knee-jerk reaction against such apparently blatant opportunism (Moffett et al. 2004), let us consider what has really been presented. The hypothesis that some women have abnormal immune function that can affect implantation is actually simply a clinical framework that can and should be rigorously tested. I would argue that this is a separate issue from the equally important and compelling drive for basic scientific understanding.

In the words of Lord Byron, ‘Time is the corrector when our judgments err’, and the reality is that NK testing is not going away (e.g., Santillan et al. 2015). A recent survey found that 69% of specialists offered 290 immune testing for couples with reproductive failure (Kwak-Kim et al. 2013). The unfortunate thing is that the politics of this debate has pushed NK studies away from mainstream reproductive medicine. This is surely not helpful for anyone. NK studies are so rarely accepted or presented at either of the main scientific meetings in reproductive medicine (European Society for Human Reproduction and Embryology and American Society for Reproductive Medicine) that it is easy to conclude that NK testing is simply not worth further evaluation. But there is no reason why mainstream clinicians should not be presented with NK studies and evaluate them just like any other clinical test and potential intervention.

In fact, recent systematic reviews are certainly not against the concept that targeted treatment of women with high NK cell activity could be beneficial. Tang et al. (2011a,b) searched through 783 publications and picked only 12 that fitted their inclusion criteria, 6 of which were suitable for final analysis. In spite of that enormous loss of data, women with unexplained reproductive failure did indeed have higher (although statistically insignificant) NK test results. This, remember, is in spite of the fact that every study used different technologies and reference ranges for assessing NK cells. More recently, Seshadri and Sunkara (2014) showed higher levels of pNK cells (but not uNK cells) levels in women with infertility and recurrent miscarriage. In one of the largest studies showing higher uNK cell numbers in women with recurrent miscarriage (Tuckerman et al. 2007), a recent re-analysis accounting for the developmental stage of the endometrium did show an association with outcome (Liu et al. 2014).

Similarly, studies of pNK cells as predictors of outcome require careful assessment. It seems that while simple enumeration of cell count is not predictive of IVF success (Thurn et al. 2005; Mattoo et al. 2007), the level of activation of pNK cells (e.g., using the CD69 activation 320 marker) is predictive (Thurn et al. 2004; Dongkai et al. 2014). In a recent review of treatment for high NK cells, Palanski et al. (2014) assessed 217 studies, of which 3 were of sufficient quality to be included in their study. Again, in spite of that enormous loss of data, immune therapy for women with high NK cells does appear to be beneficial. In 325 an RCT of 112 women undergoing IVF with high pNK cell activation, clinical pregnancy rates were 48% in the prednisolone-treated group versus 29% in the non-treatment group (Alhabsi et al. 2011). Similar benefits in live birth rates were reported in trials with intravenous immunoglobulin (Winger et al. 2011; Moraru et al. 2012). And, in spite of 330 the technological issues around uNK testing, a small RCT of women with unexplained recurrent miscarriage and high uNK cells showed similar benefits with live birth rates from prednisolone (60%) versus placebo (40%) (Tang et al. 2011a,b). It was calculated that a definitive trial would require over 850 uterine biopsies. These formal trials are summarised by other analyses (Clark et al. 2006; Li et al. 2013) and recent observational studies (Van den Heuvel et al. 2007; Hellmann et al. 2013; Krueger and Sacks 2012; Srinath et al. 2013; Kuozy and El-Rayess 2014) showing possible benefit of immune therapy targeted to those women with high NK cell activity.

The concept of targeting the right patients was illustrated in a recent paper on the use of prednisolone (Dan et al. 2015). It was shown that
prednisolone improves outcome in women with repeated miscarriage, but not in women undergoing routine (not repeated) IVF. A treatment is only likely to work if there is a problem that needs treating and that it is the right treatment. Future trials should specifically target women with recurrent miscarriage and repeated IVF failure who have high NK cell activity.

What is the harm?

Moffett and Shreve (2015) are concerned that NK testing is not only of no proven benefit but also that it is harmful. And I would completely agree that a doctor’s primary responsibility is to do no harm. But women having NK testing would certainly not consider themselves ‘normal’ or ‘healthy’ (Craig et al., 2002; Verhoek et al., 2007). In all clinical applications so far reported, NK testing is for women with major reproductive morbidity—not just one but multiple IVF failures or miscarriages. The conventional answers clinicians give to those patients (‘unexplained’, ‘unlucky’, ‘it’s a numbers game’, ‘try again’) are part of the solution, but we cannot pretend it is the only possible alternative.

The point about NK testing (if it is done well, with the technical issues discussed above taken into consideration) is that it can give patients confidence that their clinician is at least thinking about their very frustrating problem. Where is the harm in that? The current alternative is that frustration and the Internet drives patients and their doctors to randomly take immune therapy. Immune therapy for all those with unexplained reproductive failure is a well-trodden path that is not effective—or rather, not effective enough (Clark 2012; Clark 2013). Poor patient selection means more women exposed to therapies which may be of no benefit—therapies which could be detrimental, expensive or even harmful. NK testing is currently our best and most established possibility for screening for those women who are most likely to benefit from immune therapy. And while evidence for benefit is certainly not yet proven, the evidence there is benefit cannot and should not be ignored. We need more studies, and better studies. Those will be possible only with greater openness and debate, with more interaction between clinicians and researchers.

It is clear that clinicians have a heavy duty of responsibility to engage with this topic appropriately. I agree that excessive enthusiasm (along the lines described by Moffett and Shreve, 2015) may drive too many down the path of immune therapy or expensive treatments. This includes the common misconception by enthusiasts that all unexplained failures must have an immune cause. Immune testing must be robust, reliable, and, given current knowledge, always done in the context of a trial so that outcomes are monitored. But similarly, excessive negativity about immune testing shuts the door on something that ‘might’ be beneficial, and if that is the case, patients can be harmed by not getting the treatment that they need.

Moffett and Shreve (2015) describe potential harm from immune therapy. Any clinician who has used these therapies would know that the risks are very small—not negligible or zero, but very small, just like any medication or surgical procedure given to any patient for any condition. This is no different. Patients perceive a problem and, if potential risks and benefits are explained, it is a clinical decision to try immune testing and therapy. It is true that there is no currently defined ‘disease’ associated with high NK cell activity. But that is exactly the point of doing more rather than less research in this area. Perhaps, there is a genuine immune disorder, which is only manifest in failed embryo implantation and establishment of pregnancy. This is a very real hypothesis that needs to be explored.

The main areas of potential harm that I see are that (i) testing is excessive or reference ranges poorly defined leading to too many women falsely diagnosed with an immune problem and (ii) treatments are too expensive. Thus, given the current evidence and debate, rather than lower thresholds for diagnosis should be used (e.g. King et al., 2010). And, given that there is currently no evidence at all comparing the effectiveness of different immune therapy protocols, offering an expensive therapy should be a particularly cautious ethical decision. But the decision should be made by the clinician and patient because only they can judge the relative importance of such a therapy in a particular case.

Conclusions

The article that stimulated this debate paper includes a peerless description of current knowledge of the biology of NK cells. And we should certainly use that as a base for our overall understanding. But there is a problem when such knowledge is extrapolated so forcefully into the clinical arena. Patients have real problems that need pragmatic solutions rather than theories. There are many instances where scientific theory resulted in completely opposite medical outcomes to those intended, for example, the widespread use of hormone replacement therapy to reduce long-term mortality which was found, when trials were completed, to have the reverse effect. In other words, real observations should be a cue for further questioning and research in the basic science. If, for whatever reason, women with high NK cell levels benefit from immune therapy, we need to understand why. On the contrary, Moffett and Shreve (2015) propose that immune therapy and NK testing cannot possibly be of benefit according to their theory, and it should be banned or highly regulated. How can that approach possibly be of benefit to anyone, including their own research? In the words of Abraham Lincoln, “He has the right to criticize, who has the heart to help”. Controversy exists in every area of human endeavour. Immune testing in reproductive failure is no different. Over more than 20 years, in many different countries and clinics, benefit has been observed by patients and clinicians. Much of it has been presented and published, and ultimately the assessment of such studies should be made by individual clinicians rather than committees. What is needed is more widespread education on this topic, more consensus on how to perform the tests and what they indicate and more clinical trials to test the very real hypothesis that women with high NK cell activity may benefit from immune therapy. Understanding the pathological basis for this clinical observation requires a more open and two-way relationship between laboratory scientists and clinicians than Moffett and Shreve appear to be proposing.

Perhaps, the most significant advance in reproductive immunology in recent years is the understanding that the maternal immune system plays an active rather than passive role in the establishment of pregnancy (Moffett and Lake, 2004; Beaman et al., 2014; Sacks 2015). Immune testing may be useful, but we currently only have one therapeutic option—immunee suppression. More research in both clinical and laboratory areas is needed to refine testing as it is possible, and indeed probable, that some women would benefit from some form of immune stimulant. NK testing is our current most widely available and established immune test, and we need to know more about it. And we should embrace all the tools we have to make progress in this new science.
Author's roles

G.S. proposed the concept and wrote the paper.

Conflict of interest

None declared.

References


Chiaou H, Inflammation, NK cells and implantation: friend and foe (the good, the bad and the ugly); replacing placental viability in an evolutionary perspective. J Reprod Immunol 2013;97:2–13.


Fawzy M, El-Refaeiy AA. Does combined prednisolone and low molecular weight heparin have a role in unexplained implantation failure? Arch Gynecol Obstet 2014;289:677–680.


Lash GE, Laird SM, Li TC, Innen BA, Mariee N, Patel N, Querby SM, Bulmer JN. Standardisation of uterine natural killer (uNK) cell measurements in endometrium of women with recurrent miscarriage. Oral presentation at ESHRE annual meeting, Munich, Germany, 2014.


Moffett, Shreve. First do no harm: uterine NK cell in assisted reproduction.

**Q6** Human Reproduction 2015.


Sacks G, Varner N, Russell P. Natural killer cell analysis from blood, uterus, both or not at all? Oral presentation at ASPIRE, Brisbane, 2014.


