Standardisation of uterine natural killer (uNK) cell measurements in the endometrium of women with recurrent reproductive failure

LASH, Gendie E., BULMER, Judith N., LI, Tin Chiu, INNES, Barbara A., MARIEE, Najat, PATEL, Gnyaneshwari, SANDERSON, Jean, QUENBY, Siobhan and LAIRD, Susan <http://orcid.org/0000-0003-4020-9020>

Published version

LASH, Gendie E., BULMER, Judith N., LI, Tin Chiu, INNES, Barbara A., MARIEE, Najat, PATEL, Gnyaneshwari, SANDERSON, Jean, QUENBY, Siobhan and LAIRD, Susan (2016). Standardisation of uterine natural killer (uNK) cell measurements in the endometrium of women with recurrent reproductive failure. Journal of Reproductive Immunology, 116, 50-59.

Copyright and re-use policy

Standardisation of uterine natural killer (uNK) cell measurements in the endometrium of women with recurrent reproductive failure

Gendie E Lash[1], Judith N Bulmer[1] Tin Chiu Li[3,4], Barbara A Innes[1], Najat Mariee[2], Gnyaneshwari Patel[5], Jean Sanderson[6], Siobhan Quenby[5], Susan M Laird[2*]

[1]Reproductive and Vascular Biology Group, Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne

[2]Biomedical Research Centre, Sheffield Hallam University, City Campus, Sheffield

[3]Jessop Wing, Sheffield Teaching Hospitals, Sheffield

[4]Prince of Wales Hospital, Chinese University of Hong Kong, Hong Kong, China

[5]Biomedical Research Unit in Reproductive Health, Division of Reproductive Health, Warwick Medical School, University of Warwick, Warwick

[6]ScHARR, University of Sheffield, Regent Court, Sheffield.

Running title: Uterine NK cell assessment in reproductive failure

*Address for correspondence

Dr Susan Laird

Biomolecular Science Research Centre

Sheffield Hallam University

City Campus

Sheffield

S1 1WB

Email: S.M.Laird@shu.ac.uk

Phone: +44 114 225 3035

Abstract

Considerable work is being carried out on uNK cells to determine their role in pregnancy outcome. There is also debate about whether uNK cell measurements should be included in the clinical assessment for women with reproductive failure. The fact that the density of uNK cells reported by different centres varies makes advances in this field difficult. The aim of this study was to determine the reason for these differences and to develop a standardised method. Three centres exchanged five sections of endometrium from five women. Sections were immunostained for CD56. Images were taken of 10 random fields at x400 magnification; total stromal and uNK cells were counted using Image J. Results were expressed as % positive uNK cells and the variation in counts obtained in each centre was compared. After initial analysis a standardised protocol was agreed and the process repeated.

Significant variation was seen in the counts obtained after initial analysis (Centre AvsB, mean difference = -0.72 P<0.001; AvsC mean difference = -0.47 P<0.001; BvsC, mean difference = 0.25 P=0.085). Differences may be due to duration of fixation, the embedding and sectioning processes, selection of areas for assessment, definition of immunopositive cells and inclusion or exclusion of blood vessels. Adoption of a standardised protocol reduced the variation (Centre AvsB mean difference = -0.105 P=0.744; AvsC mean difference = 0.219 P=0.150; BvsC mean difference = 0.32 P=0.031). Use of a standardised method is needed to develop a meaningful clinical test for uNK cell measurements.

Key words: uterine natural killer cells; endometrium; assessment protocol; recurrent miscarriage; recurrent implantation failure

Introduction

Uterine natural killer (uNK) cells are the major leucocyte present in the endometrium at the time of implantation and early placentation. They differ phenotypically from most peripheral blood NK cells and are $CD56^{bright}$, CD16-, CD9+ with cytolytic granules and increased KIR expression (Bulmer et al. 2010 Male et al. 2011). Although peripheral blood also contains a subpopulation (10%) of phenotypically similar cells, microarray analysis has shown that the $CD56^{bright}$ CD16- NK cells in peripheral blood are distinct from uNK cells (Koopman et al. 2003). CD56+ cells are present throughout the menstrual cycle but the number increases exponentially in the mid-secretory phase starting 6 to 7 days after the LH surge, the beginning of the putative time of implantation. The number of CD56+ cells remains high during early pregnancy and comprises 70% of the lymphocytes at the interface between maternal decidua and the invading trophoblast (Bulmer & Lash 2005). The exact function of uNK cells is unclear, although their increased numbers at the time of embryo implantation and their presence adjacent to the invading trophoblast suggests that they play a role in implantation. Unlike peripheral blood NK cells, uNK cells in endometrium have lower cytolytic activity and increased cytokine production (Bulmer & Lash 2005). More recent work suggests that uNK cells produce numerous angiogenic factors and may play a role in spiral artery remodelling which is essential for establishing a successful pregnancy (Li et al. 2001; Lash et al. 2010; Robson et al. 2012).

The numbers of uNK cells are increased in the peri-implantation endometrium of a sub-population of women with recurrent miscarriage (RM) (Quenby et al. 1999; Clifford et al. 1999; Tuckerman et al. 2007) and women with recurrent implantation

failure after IVF (RIF) (Ledee-Bataille et al. 2004; Tuckerman et al. 2010), which suggests that they may play an important role in embryo implantation. However, there are contradictory reports as to whether an increased number of endometrial uNK cells correlates with pregnancy outcome in these women (Tuckerman et al 2010; Tang et al. 2011). This is in part due to a lack of understanding of the role of uNK cells in the establishment of pregnancy and therefore how their altered numbers may impact reproductive health. Despite this controversy there is an increasing demand from women with recurrent reproductive failure for an "endometrial uNK cell count test". However, even if the result shows a high number of uNK cells there are no proven successful treatments, although treatment with prednisolone has been shown to have some success in one centre (Quenby et al. 2005; Lash et al. 2011). All this has led to debate about whether measurements of uNK cells should be included in clinical assessment for women with RIF or RM (Sacks 2015; Moffett & Shreeve 2015).

One of the difficulties in advancing this area is the lack of consensus in reporting uNK cell number and in particular in defining what constitutes a "high" uNK cell count; one centre has used a cut off of 5% (Quenby et al. 2005), while another has used a cut off of 12.9% (Tuckerman et al. 2010). There is also no clear definition of a "normal range" of uNK cell numbers, partly because obtaining endometrium from normal fertile control women is difficult. In addition, for practical reasons many centres sample over a 3 day period (LH+6-LH+8 or LH+7-LH+9), but during this time period uNK cell numbers increase exponentially (Russell et al. 2011; Russell et al. 2013), which may skew clinical results dependent on the “reference range” being used.

In order to understand further the role of uNK cells in successful pregnancy outcome and to provide patients with a meaningful clinical test, a single methodological protocol for measurement of endometrial uNK cells is needed. The aim of this study was to explore why different centres report such differences in uNK cell numbers, despite apparently using the same immunostaining methodology, to develop a standardised protocol and to test whether the application of this protocol reduced the variation between centres. Until standard methodology for assessment of uNK cells is established the question of variation related to biopsy timing cannot be addressed.

Materials and methods

Study design and tissue samples

Three different centres in the UK took part in this study; Reproductive and Vascular Biology Group, Institute of Cellular Medicine, Newcastle University (Centre A); Sheffield Hallam University and Jessop Wing, Sheffield Teaching Hospitals (Centre B) and Biomedical Research Unit in Reproductive Health, Warwick University (Centre C). Local ethical committee approval was obtained for the collection of samples, and informed consent was obtained from each patient. Each centre provided 5 x 3µm formalin fixed paraffin embedded sections cut from endometrial tissue from five different patients. Samples from Centre A were collected from women undergoing hysterectomy for non-malignant conditions not affecting the endometrium and were collected at random times during the menstrual cycle. Samples from Centre B were from women with recurrent implantation failure after IVF (RIF) and were collected on LH+7-LH+9 of the cycle. Samples from Centre C were collected from women with recurrent miscarriage (RM) and were collected on days LH+6-LH+8. Each centre immunostained the 15 samples and counted the number of cells using their "usual" research or routine Pathology Department procedure as detailed below. The results obtained for each of the 15 samples in each of the centres were compared. After a review and discussion to identify reasons for the differences observed a standardised protocol was developed; the process was then repeated on a further 15 freshly collected samples (five from each centre) using the agreed protocol.

Initial analysis

The processing of tissue and staining of CD56+ cells in the initial experiment was carried out as per the established procedure in each centre detailed below (Tuckerman et al 2007; Tuckerman et al 2010; Quenby et al. 2005; Lash et al. 2012). In addition to their usual procedure Centre A assessed differences in slides stained in the routine pathology laboratory and use of a different second antibody system.

Centre A

Samples (biopsies from hysterectomy specimens) were fixed in 10% buffered formalin for 24-48 hours at room temperature prior to routine processing, embedding in paraffin wax and microtome sectioning (3μm thick sections). Immunostaining was performed within the routine Cellular Pathology laboratory (automated staining using the Ventana XT staining platform (Ventana Medical Systems, Basel, Switzerland)) and in the research laboratory (hand stained using two different detection systems, Vectastain Elite ABC kit (Vector Laboratories, Peterborough, UK) and Expose kit (AbCam, Cambridge, UK)). In the research laboratory sections were dewaxed in xylene, rehydrated through descending concentrations of alcohol to 0.15M Tris buffered 0.05M saline, pH 7.6 (TBS), incubated in 1% H_2O_2 in methanol for 10 min to block endogenous peroxidase activity and subjected to heat-mediated unmasking of antigen by pressure cooking in citrate buffer, pH 6.0 for 1 min. The sections were then incubated in anti-CD56 antibody (NCL-CD56-504; 1:100 dilution; Leica Biosystems, Newcastle upon Tyne, UK) at room temperature for 1 hour before detection with the appropriate secondary kit (as described above) and visualised using 3,3'-diaminobenzidine (DAB; Sigma Chemical Co., Poole, UK) containing 0.01% H_2O_2 to give a brown reaction product. The sections were lightly counterstained with

Mayer´s haematoxylin, dehydrated, cleared in xylene and mounted with DPX (distrene, plasticiser, xylene) synthetic resin (Raymond A. Lamb Ltd., London, UK). Negative controls included replacement of the primary antibody by appropriate non-immune serum and were performed for each antibody run. In the routine pathology laboratory staining was performed using the Ventana automated staining system (all reagents being from Ventana Medical Systems). Antigen retrieval was performed using solution CC1 for 60 min, the primary antibody used was anti-CD56 (123C3; predilute) and the detection system was Ultraview DAB kit (760-500).

Image analysis was performed by one operator who imaged 10 x400 fields using a Nikon 80i microscope and NIS Elements software (Nikon Instruments Inc., Surrey, UK). Random images were taken with the luminal epithelium in view, moving 1 field of view away between each image, and cell counts were performed using Image J (Version 1.46, NIH, Maryland, USA) cell counter plug in. The total number of stromal cells was determined by counting the number of nuclei, and the total number of uNK cells determined by counting immunopositive cells. Luminal and glandular epithelium were not included in the total stromal cell count, but blood vessels were. Data are expressed as %uNK cells/total stromal cells.

Centre B

Samples (endometrial pipelle biopsies) were fixed in 10% buffered formalin for 24-48 hours at room temperature prior to routine processing, embedding in paraffin wax and microtome sectioning (3µm thick sections). Immunostaining was performed within the research laboratory where sections were dewaxed in xylene, rehydrated through alcohols to TBS and incubated in 3% H_2O_2 in methanol for 20 min to block

endogenous peroxidase activity. Heat-mediated unmasking of antigen was performed by microwaving (800W) in citrate buffer, pH 6.0. Buffer was heated in the microwave oven until boiling, slides were added to the buffer and incubated on high for 3 min, followed by 12 min on medium and then allowed to cool for 20 min. The sections were then incubated in anti-CD56 antibody (NCL-CD56-504; 1:50 dilution; Leica Biosystems) at room temperature for 1 hour and then at 4°C overnight before detection with the Vectastain Elite ABC kit (Vector Laboratories) and DAB containing 0.01% H_2O_2 (Vector Laboratories). The sections were lightly counterstained with Mayer´s haematoxylin, dehydrated, cleared in xylene and mounted with DPX synthetic resin (Sigma Chemical Co.).

Image analysis was performed by one operator who assessed 10 x400 fields directly using an Olympus BX60 microscope (Olympus Keymed, Essex, UK). Fields of view were chosen at random throughout the tissue with the total number of stromal cells determined by counting the number of nuclei, and the total number of uNK cells determined by counting immunopositive cells. Luminal and glandular epithelium, and blood vessels were not included in the total stromal cell count. Data are expressed as %uNK cells/total stromal cells. After initial comparison of data and discussion between the operators in the 3 centres, Centre B reanalysed this initial set of slides. Images were taken with the luminal epithelium in view and cell counts were performed using Image J (Version 1.46, NIH) cell counter plug in.

Centre C

Samples (endometrial pipelle biopsies) were fixed in 10% buffered formalin for 24-120 hours at room temperature prior to routine processing, embedding in paraffin

wax and microtome sectioning (3μm thick sections). Immunostaining was performed within the routine pathology laboratory (automated staining using the Bondmax automated staining station and associated solutions (Leica Biosystems)). Antigen retrieval was performed by pressure cooking in citrate buffer, pH 6.0 for 1 min. The primary antibody used was anti-CD56 (NCL-CD56-504; 1:200 dilution) and the detection system was the Novocastra Bond Polymer Refine Detection kit.

Image analysis was performed by one operator. The slides were scanned using the MIRAX Midi Digital Slide Scanner (Carl Zeiss MicroImaging GmbH, Jena, Germany). The images were viewed using 'Panoramic Viewer' (3D HISTECH Ltd, Budapest, Hungary) and 10 x400 fields were assessed. Images were taken with the luminal epithelium in view and cell counts were performed using Image J (Version 1.46, NIH) cell counter plug in. The total number of stromal cells was determined by counting the number of nuclei, and the total number of uNK cells determined by counting immunopositive cells. Luminal and glandular epithelium were not included in the total stromal cell count, but blood vessels were. Data are expressed as %uNK cells/total stromal cells.

Standardisation analysis

Inter-observer error was determined by swapping images between Centre A and Centre B (n=13 total), each being assessed by 3 different operators (1 from Centre A and 2 from Centre B).

To determine the optimal number of total stromal cells to count, running averages assessment was performed (n=14). The %uNK cells/stromal cells in10

images/sample were determined, the mean obtained, and from this value 5% of the mean was calculated. A running average was determined by calculating the mean of images 1+2, 1+2+3, etc until the values fell within the 5% of the mean of the 10 images. The number of images and total stromal cells to reach this level of consistency was then determined (Mariee et al. 2012).

In some cases insufficient luminal epithelium was available to perform all necessary assessments at this level in the tissue. From the initial analysis and discussions we were aware that assessment of cells in the deeper endometrium gave spurious high results. We therefore aimed to determine how far from the luminal edge we could assess while still obtaining consistent results. To this end 10 samples were chosen and 4-5 x400 consecutive images from the luminal edge were obtained and assessed as described above for Centre A.

Development of an agreed protocol

Several meetings were held between the participants from the different centres to discuss the sources of variation; a standardised protocol was then developed for further testing. Using our collective experience and after comparison of each centre's original methodology it was agreed that sources of variation arose from quality of tissue fixation (including excessively long fixation), processing, image capture, selection of areas to count, and definition of immunopositive cells. To test all of these potential sources of variation individually would have required a large number of additional samples and therefore the decision was taken to alter all of these potential sources of variation. Figure 1 shows the standard procedure adopted as a result of these discussions.

The amended protocol included fixation of the tissue in neutral buffered formalin for 24-48 hours at room temperature and ensuring that water baths used during sectioning were kept dust free. The immunostaining was carried out as described above for each centre, with Centre A using the Vectastain Elite ABC kit in the research laboratory. Images for analysis were selected adjacent to the luminal edge, or as close as possible within 5 fields, and captured digitally. Cells were counted using the cell counter plug in in Image J, with at least 3800 total stromal cells counted. All stromal cells were counted, including endothelial cells and smooth muscle cells of the blood vessels, but excluding glandular and luminal epithelium. The % uNK cells/total stromal cells for each image was calculated; the final cell count was reported as the mean of all counted images.

Statistical analysis

Data are presented as means $\pm$ SD. Agreement between the counts obtained by each centre, by each of the three different methods within one centre and between single operators was assessed using a linear mixed effects model (Roy 2009). Since the counts were not normally distributed the logit-transform of the proportion of positive cell counts were used for the analysis. Analyses were conducted using the nlme package in the R statistical software package. $P<0.05$ was considered significantly different.

Results

Variation in uNK cell numbers after initial assessment

One of the samples provided by Centre A for the initial analysis did not contain a luminal epithelial border and therefore was not included in the analysis, leaving n=14 for this part of the study. Mean (± SD) of cell counts for each sample assessed under all conditions by each of the different centres is shown in Table 1.

There was considerable variation in the % CD56+/total stromal cells within the same endometrial biopsy sample reported from each centre after the initial analysis (Table 1 and Figure 2A). This was explained, at least in part, by the fact that one of the centres (Centre B) did not use Image J for cell counting, assessing slides manually and also did not always count cells adjacent to the luminal epithelial edge. This centre then re-analysed the slides that they had stained using Image J making sure that a luminal edge was included in all fields. This reduced, but did not eliminate the variation (Figure 2B). Significant variation was still seen in the counts obtained (Centre A vs. B, mean difference = -0.72 $P<0.001$; A vs. C mean difference = - 0.47 $P<0.001$; B vs. C, mean difference = 0.25 $P=0.085$).

Figure 3 illustrates the differences in immunostaining patterns obtained from each centre on slides provided by the three different centres. Tissue processed in Centre A showed clear positive immunostaining around the periphery of the cells, while samples processed in the other two centres showed additional specks of positive immunostaining that were not associated with cell nuclei. These observations were irrespective of which centre performed the immunostaining and therefore suggested an issue in the fixation, processing and section preparation processes.

Variation between staining methods and assessment operator

The variation in % CD56+ cells obtained in Centre A when the analysis was carried out in the routine pathology laboratory and in the research setting using two different secondary antibody systems (Vectorstain Elite ABC or the Expose kit) was also assessed. Although there were differences in the number of positive cells reported, the variation was considerably less than the variation between the three centres (Expose vs. ABC, mean difference = -0.025 P=0.678; Expose vs. Routine Pathology, mean difference = 0.15 P=0.026; ABC vs. Routine, mean difference = -0.18 P=0.004) (Figure 4A).

There was also some evidence for variation in the counts obtained when the same image was counted by 3 different observers (n=13), although the magnitude was not as large as between centre variation (Operator 1 vs. Operator 2 mean difference = 0.355 p=0.019; Operator 1 vs. Operator 3 mean difference = 0.089 P=0.315; Operator 2 vs. Operator 3 mean difference = -0.266 P = 0.0265) (Figure 4B).

Numbers of fields/cells to be counted

The distribution of uNK cells within the tissue is not uniform. In addition despite all sections being photographed at x400 magnification, the images appeared to show different degrees of magnification (Figure 3). This was due to different microscope camera systems with different magnifications or 'camera factors'. To determine the total number of cells that need to be counted to overcome both these issues we used the running average method (Mariee et al. 2012) to determine that 3800 stromal cells needed to be counted to provide a true and reliable cell count (Figure 4C).

Depth of field for analysis

Not all samples contained sufficient luminal epithelium for 10 adjacent x400 fields or 3800 total stromal cells to be included. Given that samples from patients suffering reproductive failure are likely to be pipelle biopsies this could present a problem in the clinical situation. To determine the effect of depth from luminal surface on the fields chosen for counting, individual x400 fields were chosen up to 5 x400 fields from the endometrial luminal edge as shown in Figure 5A. The number of cells was similar when fields were chosen up to 4 fields away from the luminal edge (Figure 5B).

Implementation of standardised protocol

Fifteen new samples (5 from each centre) were processed, stained and counted using the agreed protocol (Figure 1). Two of the samples provided by Centre A did not contain a luminal epithelial border and were not included in the analysis, leaving n=13 for this part of the study. Mean (± SD) for all samples assessed in each of the different centres is shown in Table 2. Reduced variation was seen in the counts obtained (Centre A vs. B mean difference = -0.105 P=0.744; A vs. C mean difference = 0.219 P=0.150; B vs. C mean difference = 0.32 P=0.031 (Figure 6A). In addition, images taken in each centre of the different centre's samples were more similar to each other (Figure 6B) compared with the initial assessment (Figure 3).

Discussion and Conclusion

There is considerable evidence that uNK cells may play a role in successful embryo implantation, formation of the placenta, control of trophoblast invasion and spiral artery transformation and therefore pregnancy outcome (Lash et al. 2010; Moffatt-King 2002). However, advancement in our understanding of this mechanism and how measurement of endometrial uNK cell numbers may be used clinically is hindered by a lack of consistency between numbers of cells reported between centres and a clear definition of the 'normal' range for endometrial uNK cells. The normal range of endometrial NK cells needs to be determined before its clinical application to patients with poor reproductive outcomes; sampling of endometrium from normal fertile control women is possible, but if methodologies result in different reported values from site to site the normal range cannot be established.

In this study we have investigated why three different centres in the UK report very different endometrial NK cell numbers, despite apparently using the same immunohistochemical method to identify positive cells. Each centre stained five sections provided by each of the three centres and counted the cells according to their original protocol. The results were discussed and a strict protocol developed to try and eliminate variation. Sample processing, staining and counting was then repeated taking into account the new protocol. The variation in the counts reported for each sample was reduced substantially. We were not able to change each parameter individually and therefore cannot pinpoint the exact source of the variation in the original analysis, but the difference in appearance of sections prepared from tissue collected and processed in the 3 different centres suggest that it may be due to duration of tissue fixation, differences in the wax embedding or sectioning

processes. To eliminate the variation in tissue processing the samples were fixed in 10% neutral buffered formalin for 24-48 hours at room temperature and processed conventionally with xylene.

One of the factors identified as contributing to the variation was the fact that one centre was not counting cells adjacent to the luminal edge, but instead was choosing fields at random across the tissue. The distribution of leucocytes within the endometrium is not uniform; there is clustering of leucocytes, including CD56+ cells, particularly around glands and blood vessels (Bulmer & Lash 2005). This is also the reason for the sometimes quite considerable variation in the counts obtained for each field in an individual sample. To obtain consistent results it was calculated that at least 3800 stromal cells need to be counted; this agrees with previously published work (Mariee et al. 2012). Expressing this parameter as a cell number rather than a number of fields is necessary to overcome the camera factors associated with different microscope and digital camera systems.

The method used to select fields for counting was discussed extensively. Previously the fields for counting were chosen ‘at random’ across the tissue. However, to be truly ‘random’ in selection of fields for quantification is extremely difficult and there is a tendency to choose fields where there are positive cells and ignore fields where all cells are negative. In order to prevent observers only counting fields that contained positive cells the first field to be captured was selected at random, ensuring that it contained the luminal epithelial border. Subsequent fields were obtained by moving one field to the left or right of the original field (skipping a field between each captured image), keeping the luminal epithelial border in view and repeated until 10

fields had been captured. If there were not 10 fields containing the luminal epithelium in view, we have demonstrated that using up to 4 consecutive fields extending deeper into the tissue will still provide valid results.

Variation also arose from definition of an immunopositive cell; with some observers counting brown staining not associated with a cell nucleus, it was therefore recommended that a cell nucleus must be visualised with the immunopositive membrane staining to assess a cell as positive. Another issue was the counting of cells around blood vessels, with some assessors not including this compartment in their cell counts. However, it is not always easy to distinguish these cells from the stromal population, especially if the vessel has poorly developed muscle layers. For a protocol to be successful it needs to simple and easy to use and therefore it was decided that cells across the whole of the stromal compartment should be included, although luminal and glandular epithelial cells, which are easily distinguished are excluded.

One of the limitations of the study is the sample size (five samples from 3 different centres - 15 samples in total). However, the aim of this study was simply to determine why reports of uNK cell counts are so different and to identify the key steps or factors which are important in preventing this happening in the future. This sample size was large enough to achieve this and will enable further work, using larger numbers to take these factors into account.

Numbers of uNK cells increase exponentially as the menstrual cycle progresses (Bulmer & Lash 2005; Russell et al. 2011; Russell et al. 2013) and therefore timing of

the biopsy is critical or a reference range of ‘normal’ uNK cell numbers throughout the menstrual cycle needs to be established. However, before this can be addressed a method which provides consistency between different centres is required and this study has addressed this. Further work is required to determine what constitutes a ‘high’ uNK cell density and the timing of the biopsy will be an important aspect of this work. In addition, it still needs to be determined whether the increased uNK cell numbers reflects a causative role in recurrent miscarriage and recurrent implantation failure or rather is a marker of a more generalised endometrial dysfunction that contributes to these conditions (Salker et al. 2010; Teklenburg et al. 2010).

Conclusion

This study shows the importance of bringing quality control processes into the measurement of uNK cells and the need to establish a quality control methodology prior to extension to clinical practice. The method described in this paper may not be perfect, but it is producing consistent results and will enable comparison of results between centres. A larger sample size and the inclusion of different laboratories is now required to fully refine and validate this protocol; of particular interest will be further definition of the field of view selected for assessment. In addition, standardized collection and assessment of samples from normal fertile control women will enable us to establish a "normal" range for endometrial uNK cells and to determine their role in implantation.

Study funding/competing interests: The study was not supported by external funding sources. The authors have no competing interests to declare.

Table 1. Initial assessment (Mean ± SD)

	Centre A – Expose kit	Centre A – ABC Vector kit	Centre A – Routine Pathology	Centre B – Manual	Centre B – Image J	Centre C
A	4.7 ± 1.5	3.9 ± 1.4	2.3 ± 0.8	24.9 ± 6.7	12.2 ±2.7	12.6 ± 1.6
B	5.8 ± 2.3	4.4 ± 1.3	4.1 ± 0.9	27.4 ± 5.9	14.0 ± 5.6	9.8 ± 0.5
C	2.2 ± 0.9	3.3 ± 1.6	2.0 ± 0.9	26.4 ± 12.5	6.2 ± 3.1	4.5 ± 0.2
D	3.2 ± 0.2	3.7 ± 1.2	3.6 ± 0.2	8.9 ± 2.4	6.9 ± 1.8	4.5 ± 0.8
E	5.1 ± 2.2	5.2 ± 2.5	5.8 ± 2.5	13.1 ± 6.2	24.6 ± 10.2	7.7 ± 0.8
F	2.0 ± 0.7	2.7 ± 1.4	2.0 ± 1.1	4.7 ± 2.2	1.9 ± 1.2	3.2 ± 0.06
G	9.8 ± 2.9	9.1 ± 2.6	8.9 ± 2.1	18.0 ± 5.1	20.1 ± 7.9	13.1 ± 0.9
H	4.3 ± 2.1	4.4 ± 1.6	3.8 ± 1.0	10.5 ± 4.6	8.7 ± 3.8	4.4 ± 0.8
I	6.8 ± 1.4	9.8 ± 2.0	7.0 ± 2.7	22.0 ± 9.1	15.3 ± 5.4	17.4 ± 1.1
J	5.3±1.9	5.3±2.2	5.3±1.7	14±2.8	10.6±5.3	10.2±0.8
K	2.0±1.6	2.1+0.6	1.8±0.9	5.7±1.7	2.5±0.8	2.8±0.2
L	27.8 ± 5.3	30.8 ± 5.7	25.4 ± 3.6	37.0 ± 4.7	38.8 ± 5.3	24.5 ± 3.9
M	1.5 ± 0.6	1.5 ± 0.8	1.7 ± 0.4	6.5 ± 1.4	3.0 ± 1.7	2.2 ± 0.1
N	1.7 ± 0.5	1.3 ± 0.5	0.9 ± 0.5	3.7 ± 2.0	2.5 ± 1.5	1.8 ± 0.07

475 **Table 2. Validation assessment (Mean + SD)**

	Centre A	Centre B	Centre C
A	4.8 ± 9	4.5 ± 0.7	3.7 ± 0.7
B	6.5 ± 0.7	3.4 ± 0.3	4.6 ± 0.3
C	2.9 ± 0.6	4.5 ± 0.4	3.8 ± 0.3
D	6.5 ± 0.6	3.5 ± 0.5	3.3 ± 0.2
E	5.7 ± 0.5	2.2 ± 0.2	1.8 ± 0.3
F	9.8 ± 0.7	6.6 ± 0.6	7.2 ± 0.9
G	2.4 ± 0.4	3.0 ± 0.9	1.4 ± 0.2
H	1.7 ± 0.4	1.1 ± 0.3	1.0 ± 0.3
I	13.1 ± 1	13.3 ± 1.4	15.1± 0.6
J	2.7 ± 0.4	2.2 ± 0.5	2.0 ± 0.2
K	16.7 ± 2.2	22.7 ± 1.8	12.6 ± 1.8
L	2.5 ± 0.2	3.8 ± 0.6	3.8 ± 0.2
M	1.4 ± 0.3	2.2 ± 0.5	0.6 ± 0.1

References

Bulmer JN and Lash GE. Human uterine natural killer cells: a reappraisal. *Mol Immunol* 2005;**42**:511-21.

Bulmer JN, Williams PJ and Lash GE. Immune cells in the placental bed. *Int J Dev Biol*. 2010;**54:**281-294.

Clifford K, Flanagan AM and Regan L. Endometrial CD26+ natural killer cells in women with recurrent miscarriage: a histomorphometric study. *Hum Reprod* 1999;**14**:2727-2730.

Koopman LA, Kopcow HD, Rybalov B, Boyson JE, Orange JS, Schatz F, Masch R, Lockwood CJ, Schachter AD, Park PJ and Strominger JL. Human decidual natural killer cells are a unique NK cell subset with immunomodulatory potential. *J Exp Med* 2003;**198**:1201-1212.

Male V, Sharkey A, Masters L, Kennedy PR, Farrell LE and Moffett A. The effect of pregnancy on the uterine NK cell KIR repertoire. *Eur J Immunol 2011*;**41**:3017-3027.

Lash GE, Robson SC and Bulmer JN. Functional role of uterine natural killer (uNK) cells in human early pregnancy decidua. *Placenta* 2010;**31**:S87-92.

Lash GE, Bulmer JN, Innes BA, Drury JA, Robson SC and Quenby S. Prednisolone treatment reduces endometrial spiral artery development in women with recurrent miscarriage. *Angiogenesis* 2011;**14**:523-32.

Lash GE, Innes BA, Drury JA, Robson SC, Quenby S and Bulmer JN. Localization of angiogenic growth factors and their receptors in the human endometrium throughout the menstrual cycle and in recurrent miscarriage. *HumReprod* 2012;**27**:183-95.

Ledee-Bataille N, Dubanchet S, Coulomb-L'hermone A, Durand-Gasselin I, Frydman R and Chaouat G. A new role for natural killer cells, interleukin (IL)- 12 and IL-18 in repeated implantation failure after in vitro fertilization. *Fertil Steril* 2004;**81**:59-65.

Li XF, Charnock-Jones DS, Zhang E, Hiby S, Malik S, Day K, Licence D, Bowen JM, Gardner L, King A, Loke YW and Smith SK. Angiogenic growth factor messenger ribonucleic acids in uterine natural killer cells. *J Clin Endocrinol Metab* 2001;**86**:1823-34.

Mariee N, Tuckerman E, Ali A, Li W, Laird S and Li TC. The observer and cycle-to-cycle variability in the measurement of uterine natural killer cells by immunohistochemistry. *J Reprod Immunol* 2012;**95**:93-100.

Moffett-King A. Natural killer cells and pregnancy. *Nat Rev Immunol* 2002; **2**:656-663.

Moffett A, Shreeve N. First do no harm: uterine natural killer cells in assisted reproduction. *Hum Reprod* 2015; **30**: 1519-1525.

Quenby S, Bates M, Doig T, Brewster J, Lewis-Jones DI, Johnson PM and Vince G. Pre-implantation endometrial leukocytes in women with recurrent miscarriage. *Hum Reprod* 1999;**14**:2386-91.

Quenby S, Kalumbi C, Bates M, Farquharson R, and Vince G. Prednisolone reduces pre-conceptual natural killer cells in women with recurrent miscarriage. *Fertil Steril* 2005;**84**:980-4.

Robson A, Harris LK, Innes BA, Lash GE, Aljunaidy MM, Aplin JD, Baker PN, Robson SC and Bulmer JN. Uterine natural killer cells initiate spiral artery remodelling in human pregnancy. *The FASEB J* 2012;**26**:4876-85.

Roy A. An application of linear mixed effects model to assess the agreement between two methods with replicated observations. *J Biopharmaceutical Statistics* 2009;**19**:150-173.

Russell P, Anderson L, Lieberman D, Tremellen K, Yilmaz H, Cheerala B and Sacks G. The distribution of immune cells and macrophages in the endometrium of women with recurrent reproductive failure I: Techniques. *J Reprod Immunol* 2011;**91**:90-102.

Russell P, Sacks G, Tremellen K and Gee A. The distribution of immune cells and macrophages in the endometrium of women with recurrent reproductive failure. III: Further observations and reference ranges. *Pathology* 2013;**45**:393-401.

Sacks G. Enough! Stop the arguments and get on with the science of natural killer cell testing. *Hum Reprod* 2015; **30**: 1526-1531.

Salker M, Teklenburg G, Molokhia M, Lavery S, Trew G, Aojanepong T, Mardon HJ, Lokugamage AU, Rai R, Landles C, Roelen BA, Quenby S, Kuijk EW, Kavelaars A, Heijnen CJ, Regan L, Macklon NS and Brosens JJ. Natural selection of human embryos: impaired decidualization of endometrium disables embryo-maternal interactions and causes recurrent pregnancy loss. *PLoS One* 2010;**5**:e10287.

Tang AW, Alfirevic Z and Quenby S. Natural killer cells and pregnancy outcome in women with recurrent miscarriage and infertility; a systematic review. *Hum Reprod* 2011;**26**:1971-80.

Teklenburg G, Salker M, Molokhia M, Lavery S, Trew G, Aojanepong T, Mardon HJ, Lokugamage AU, Rai R, Landles C, Roelen BA, Quenby S, Kuijk EW, Kavelaars A, Heijnen CJ, Regan L, Brosens JJ and Macklon NS. Natural selection of human embryos: decidualizing endometrial stromal cells serve as sensors of embryo quality upon implantation. *PLoS One* 2010;**5**:e10258.

Tuckerman E, Laird SM, Prakash A and Li TC. Prognostic value of the measurement of uterine natural killer cells in the endometrium of women with recurrent miscarriage. *Hum Reprod* 2007;**22**:2208-13.

550 Tuckerman E, Mariee N, Prakash A, Li TC and Laird S. Uterine natural killer cells in peri-implantation endometrium from women with repeated implantation failure after IVF. *J Reprod Immunol* 2010;**87**:60-6.

Figure legends

Figure 1: Schematic of the standardised protocol.

Figure 2: Comparison of % CD56+ cells reported for the same samples from each of the different centres. **A)** Initial analysis. **B)** After re-analysis by Centre B using Image J and inclusion of the luminal edge.

Figure 3: Photomicrographs showing differences in staining obtained when staining was carried out in one centre on samples processed at the three different centres **(A)**. **B)** Higher magnification photomicrographs to demonstrate immunopositive uNK cells (left) and specks of brown DAB reactivity not fully associated with a cell nucleus (right).

Figure 4: Comparison of % CD56+ cells in the different samples when **A)** analysed by three different methods (Expose polymer based kit, Abcam; ABC Vector kit, Vector laboratories; Routine Cellular Pathology laboratory, Ventana Medical Systems) in the same centre, and counted by the same operator and **B)** three different operators assessed the same images. **C)** Representative graph of running averages to determine the total number of stromal cells to count to achieve consistent results.

Figure 5: The effect of counting cells in fields further away from the endometrial luminal edge. **A)** Illustrates how the fields can be chosen in one particular sample. **B)** shows the % CD56+ cells in each field up to 4 or 5 deep to the luminal edge (n=10 different samples).

Figure 6: **A)** Comparison of % CD56+ cells reported for the same samples from each of the different centres after adopting the agreed protocol. **B)** Photomicrographs showing reduced visible differences (compared with Figure 3) in

staining obtained when staining was carried out in one centre on samples processed at the three different centres.