HABSelect

Statistical Analysis Plan

Version: 1.0
Date: 01/02/2017

Person(s) contributing to the analysis plan	
Name(s) and position(s)	Professor Robert M West – Mechanistic statistician
Authorisation	
Position	**Chief or principal investigator**
Name	Dr David Miller
Signature	
Date	01/02/2017
Position	**Statistician for mechanistic study**
Name	Professor Robert M West
Signature	
Date	01/02/2017
Position	**Independent statistician***
Name	
Tick once reviewed	
Date	01/02/2017

*This will be the Trial Steering Committee (TSC) Statistician

Abstract

Background: *Intra-Cytoplasmic Sperm Injection is a procedure in Assisted Reproduction Technologies, where, instead of the egg being the final arbiter for selection, the 'right' single sperm is selected for each egg by the embryologist. Polyvinylpyrrolidone is normally used to slow down sperm sufficiently for capture, but clinically relevant studies suggest that using a hyaluronic acid based selection procedure (using hyaluronic acid coated plates) increases live birth rates as well as having a number of other positive effects. The HAB Select mechanistic study aims to explore how the role played by hyaluronic acid in the ICSI procedure.*

Methods/Design: *The first component of the study examines the relationships between measures of sperm DNA integrity obtained in the study and with patient demographics. Tests of DNA integrity include Comet (REF), TUNEL (REF), HALO (REF), Acridine Orange (slide based) and Aniline Blue (REF). Patient demographics will include age, sperm concentration and hyaluronic acid (HA) binding score (HBS). The second and main component aims to explore the relationships between clinical and mechanistic outcomes using a combination of decision tree analysis and linear regression modeling.*

Conclusion: *The HAB Select mechanistic study will relate clinical outcomes to sperm DNA quality.*

Trial Registration: *HABSelect is registered in ISRCTN under ISRCTN99214271*

Keywords: *HAB Select Trial; Intra-Cytoplasmic Sperm Injection; Live Birth Rate; Polyvinylpyrrolidone; Hyaluronic Acid; Hyaluronan; Sperm Selection; Assisted Reproduction Technologies; Clinical Pregnancy; Miscarriage; Male Fertility; Randomised Controlled Trial; Statistical Analysis Plan*

Introduction

In 2008, almost 40,000 couples in the UK alone were treated with assisted reproduction technologies (ART), compromising of 50,687 in vitro fertilisation (IVF) cycles. This number is set to rise in the coming years (20). Currently live birth rates for ART are at an average of 24% per treatment cycle, although live birth rates per couple are higher at 32%, because couples normally receive an average of approximately 1.3 treatment cycles. While it is estimated that more than two thirds of naturally conceived pregnancies end in failure, the limit for improvements in live birth rates following ART may not have been reached.

For all ART procedures including intra-cytoplasmic sperm injection (ICSI), the embryologist seeks to use the best sperm available. Selection is aided by sperm 'washing' techniques including density gradient centrifugation (DGC) that can enrich for sperm with high motility and good morphology ((WHO Manual, 2010) (2)). In contrast with standard IVF, where the egg is the final arbiter for selection, ICSI is dependent on the relatively subjective judgment of the embryologist to choose the 'right' single sperm for each egg.

Various studies have shown clear inverse relationships between DNA damaged sperm in the ejaculate and clinical pregnancy or live birth rates in standard IVF, but this relationship is less obvious with ICSI cycles (3). We recently reported reductions in levels of sperm DNA fragmentation following density gradient washing of semen and while the values from washed semen were reduced, they were still over twice as high in the non-pregnant (approximately 50%) versus pregnant (approximately 23%) cohorts. These and other data suggest that sperm with poor DNA quality persist in washed sperm preparations from fertile and infertile men (4-13) and unlike IVF, where there is a natural selection by the egg, ICSI could be particularly vulnerable to a poor choice of sperm. By eliminating abnormal sperm from the sample preparation for ICSI, success rates should rise accordingly.

It has been shown that immature sperm with higher rates of DNA damage have a dysfunctional ability to bind to hyaluronic acid (14, 15), which is the major glycosaminoglycan secretion of the cervix. In many clinics, polyvinylpyrrolidone (PVP) is normally used to slow sperm down sufficiently for capture in ICSI procedures (PVP-ICSI). However, clinically relevant studies (16, 17) suggest that using hyaluronic acid-selected (using hyaluronic acid coated plates) intra-cytoplasmic sperm injection (PICSI) instead of PVP-ICSI increases live birth rate, numbers of grade one embryos and clinical pregnancies, as well decreasing miscarriage rate. There is also strong evidence that PICSI reduces early pregnancy failure (17).

The HAB Select (Hyaluronic Acid Binding Sperm Selection) trial aims to confirm this by comparing the use of PICSI to PVP-ICSI procedures for the treatment of male fertility in a rigorous randomised controlled efficacy trial. Alongside the trial, a mechanistic study will be conducted. Excess material from ejaculate used in treatment

will be processed as samples, undergoing a number of laboratory tests to ascertain DNA fragmentation and chromatin compaction. Although there is independent validation of these test procedures, it is not yet clear which tests provide the best measurements or how these tests might be combined to achieve better measurements.

The main purpose of the mechanistic study is to investigate the relationships between DNA fragmentation, chromatin compaction, HA binding, and pregnancy outcomes.

This paper describes the statistical analysis plan (SAP) for mechanistic evaluation of the action of hyaluronic sperm selection.

Study Overview and Design

Overview: The HAB Select mechanistic study extends the investigation of the HAB Select trial by employing structural equation modelling (27) to define sperm integrity, descriptive summary statistics stratified by outcome and the classification tools of logistic regression and classification trees.

Study Population: The study population represents couples undergoing ICSI procedure, with the ability to provide informed consent. The following inclusion criteria are also imposed:

Women:

- BMI: 19.0 – 35.0 kg/m^2
- FSH level: 3.0 – 20.0 miU/ml and / or AMH ≥ 1.5 pmol/L
- Age: 18-43

Men:

- Age: 18 – 55
- Able to produce freshly ejaculated sperm for the treatment cycle

The exclusion criteria for the trial are as follows:

- Couples who have not consented prior to ICSI will be ineligible
- Couples using non-ejaculated sperm
- Couples using donor gametes
- Men with vasectomy reversal; cancer treatment involving any chemotherapy and / or radiotherapy in the past two years
- Previous participation in the HABSelect trial
- Split IVF / ICSI procedures

- If both FSH and AMH are tested and either of them falls outside the accepted range

There are 16 planned participating centres. Recruitment rates will be monitored and optional additional centres may be added as required. Centres will be IVF licensed hospitals and must be able to provide appointments in a dedicated clinic in which to see participants.

Consent: Written informed consent will be obtained by the principal investigator, or by another suitably qualified member of the trial team. This will compromise of a written consent form, and will be obtained for each couple before enrolment in the trial. Patients have the right to refuse consent and / or withdraw from the study at any time without giving reasons and without prejudicing any further treatment.

Treatment Procedures: In the non-interventional arm (PVP-ICSI) density gradient washed and prepared motile sperm are selected for ICSI by adding the sperm suspension to PVP on an inverted microscope. Sperm motility is slowed sufficiently to allow capture by the experience embryologist who then immobilises the sperm by crushing its flagellum with the injection pipette. The sperm is then taken up into the injection pipette and injected directly into the egg. In the interventional arm (PICSI) exactly the same procedure is carried out except that the washed and prepared motile sperm are allowed to interact with and immobilised sperm picked up from the PICSI substrate (the PISCI plates are used) beforehand.

A diagram showing the flow of samples into the study for the purpose of DNA integrity analysis is presented in Figure 1. The main aims of the mechanistic study are to gain more understanding of the relationships between hyaluronic acid binding, sperm DNA integrity (sperm quality) and the clinical performance of PICSI/ICSI.

Primary outcome.

Relationships between clinical outcomes and tests of sperm DNA integrity will be assessed by statistical modelling. Results will be reported solely for the purpose of hypothesis generation.

Secondary outcomes:

Relationships between assays measuring DNA integrity and alternative measures of sperm viablity including HBS, sperm concentration and motility will also be explored, again for the purpose of hypothesis generation.

Primary outcome measure: Live birth at ≥ 37 weeks gestation.

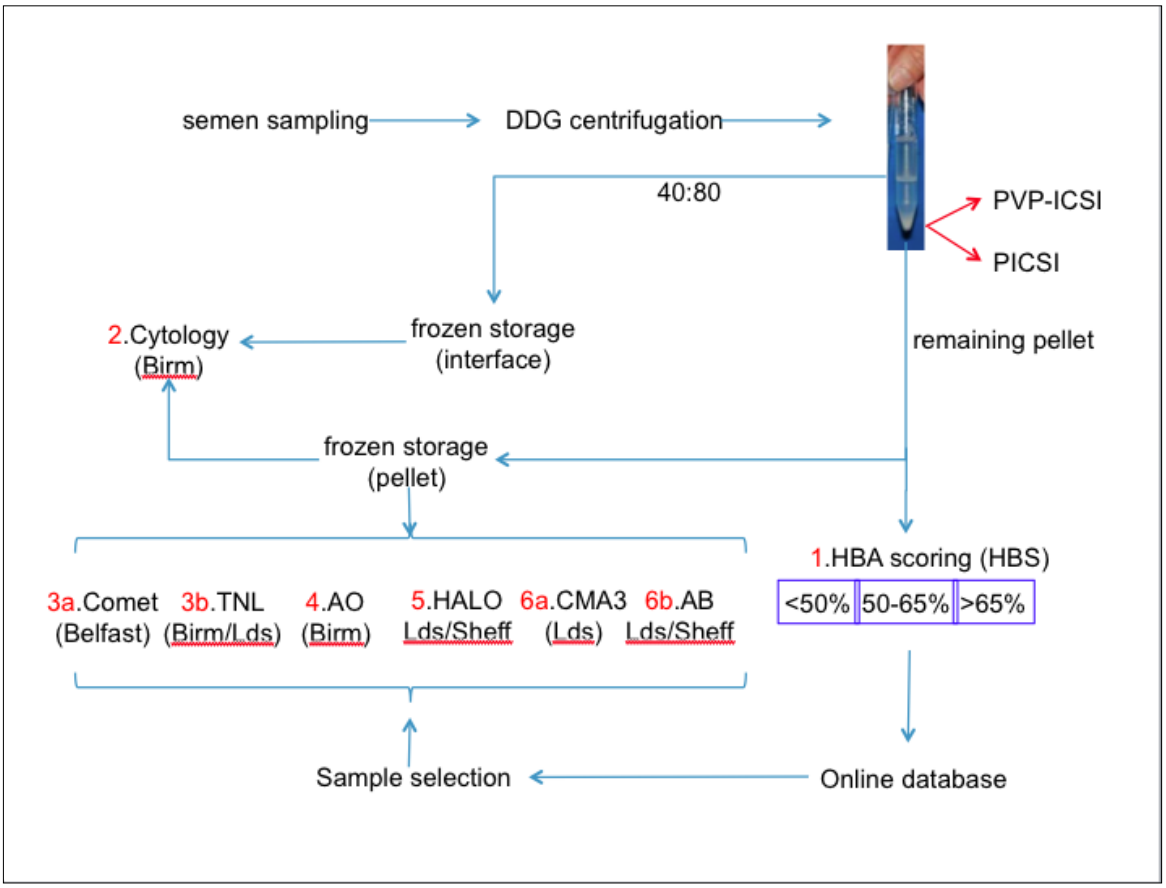


Figure 2. Schematic of mechanistic sampling. There is a hierarchical priority of testing, with the number carried out depending on the number (Table 2, **N**) of available sample aliquots after HBA[1] scoring. The testing priority is 1, 2, 3a AND/OR 3b, 4, 5, 6a AND/OR 6b. If four aliquots are available, all tests can be carried out across the four centres. If three aliquots are available, going to three centres, it will not be possible to replicate tests 5 and 6 in both Lds and Sheff. If two aliquots are available, going to two centres the same restrictions will apply and additionally, only one of tests 4, 5 or 6 will be possible. All samples will be stored centrally in Birmingham HBRC; hence they can begin to process and assess samples for distribution as soon as they are ready. If following cytology (1), there is only sufficient sample for analysis at one centre, it will be Birm who carries out any additional tests. PCTU will assist in maximising the use of available samples by all mechanistic labs. Collection of sperm from the 40:80 interface will occasionally be undertaken on CI request for the mechanistic studies.

General Statistical Considerations

This study is hypothesis generating and does not intend to test hypotheses. The purpose is to explore relationships between clinical and mechanistic measures/outcomes through classification tree analysis with logistic regression analysis. The aim is to gain a greater understanding of the potential mechanisms behind the clinical outcomes of the HAB Select trial.

Missing data: There is a threat to bias from missing data. It is plausible that the availability of sufficient sample residue to undertake the laboratory tests is related to hyaluronic acid binding or to sperm integrity. The analysis will proceed with complete cases only, and the potential bias declared as a limitation together with a possible direction of bias.

Statistical Analysis

Descriptive Statistics

Analysis will commence with descriptive statistics, reporting means and standard deviations for continuous variables and counts for categorical variables. Tables of descriptives shall be stratified by categorical outcomes. For example when the analysis considers clinical pregnancy, characteristics of couples shall be summaries for those achieving clinical pregnancy and those not. For ease of comparison, *t*-tests will be performed for continuous variables and Pearson's chi-square test for categorical variables. All analyses will be undertaken using R statistical software (29).

Structural Equation Modelling

A path diagram representing one plausible view of relationships is presented in Figure 2. Note that this shows a measurement model for DNA fragmentation where fragmentation is measured by the laboratory tests Comet, Tunel, and Acridine Orange. An important further laboratory test which reflects DNA fragmentation is HALO which also contributes to both the measurement model for DNA fragmentation and chromatin compaction. It is anticipated that the two latent variables DNA fragmentation and chromatin compaction will be strongly correlated.

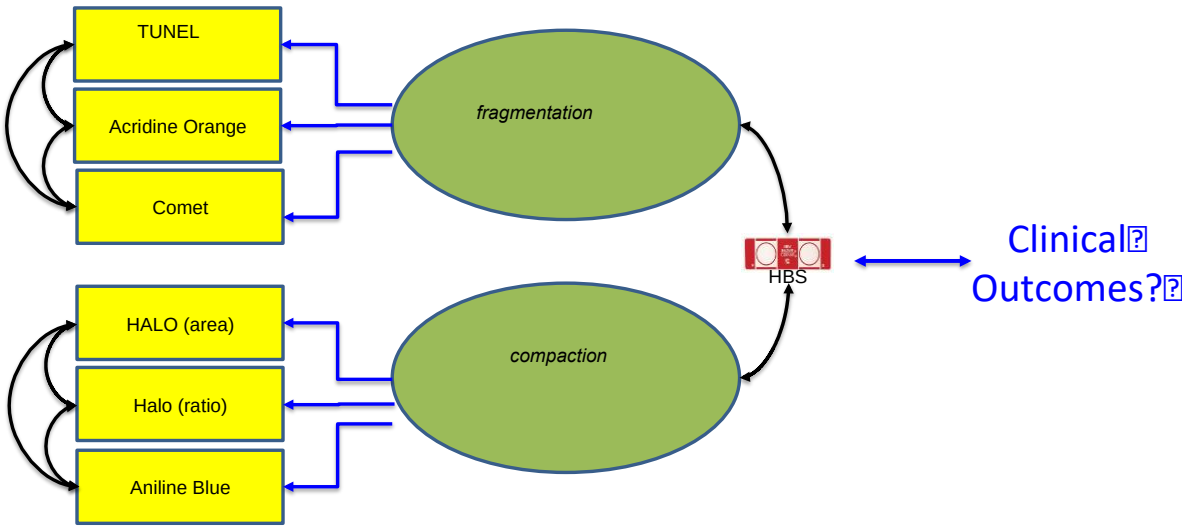


Figure 2 Structural Equation Modelling showing relationships between measured quantities (boxes) and latent variables (ellipses). *Fragmentation* is the latent variable for DNA fragmentation and *compaction* is the latent variable for chromatin compaction. In the model, Halo (area and ratio) was associated with the *compaction* variable.

It is anticipated that the association between hyaluronic acid binding and live birth rate may be nonlinear and that the greatest benefit of sperm selection through hyaluronic binding will occur when binding is poor. The path diagram does not include the influence of either treatment (PICSI rather than ICSI), the Nelson–Lawler log odds or other factors. Each of these last two covariates will influence the four pregnancy clinical outcome variables: biochemical pregnancy, clinical pregnancy, live birth and miscarriage.biochemical pregnancy, clinical pregnancy, live birth and miscarriage.

.

Sensitivity Analyses: If there is evidence that the clinical pregnancy rate differs between the treatment arms, then as a sensitivity analysis, the analysis will be redone taking only women who experience a clinical pregnancy as the denominator.

As an additional sensitivity analyses, the analysis will be redone adjusting for other factors believed to potentially prognostic or associated with the outcome, Factors to be included in this additional analysis are

- Female partner BMI (Normal: 19≤BMI<25, Overweight: 25<BMI≤30, Obese: 30≤BMI≤35)
- Female partner ethnicity (White / Asian or Asian British / Black or Black British / Other)
- Female and male age
- History of previous pregnancy (Yes/No)

In all cases, results of the primary analysis will be given more weight than those of any secondary analyses.

Subgroup Analysis: The following subgroup analyses will be performed for the primary outcome:

- Analysis of treatment effect by HBS (high (>65%) versus low (≤65%))
- Analysis of treatment effect by maternal age (≤35 years verses >35)
- Analysis of treatment effect by number of previous miscarriages (0 versus >0)
- Analysis of treatment effect by Follicle stimulating hormone (FSH) hormone level (<6.0miU/ml versus ≥ 6.0miU/ml) or Anti-Mullerian Hormone (AMH) hormone level (<17pmol/L versus ≥ pmol/L) where FSH testing is not available
- Analysis of treatment effect by sperm concentration (<15mml versus ≥ 15mml)
- We may also analyse treatment effect by a very low HBS sub-group (≤25%) versus a low HBS sub-group (>25%, ≤65%)

.

Secondary Analysis: The proportions of each secondary outcome will be compared between treatment arms using univariable and multivariable logistic regression models to estimated crude and adjusted odds ratios. A 95% confidence interval and two sided

P-value will be reported in each case. The adjusted analysis will adjust for the same factors as the primary analysis.

To enable further exploration of effects and as an alternative analysis to logistic regression, classification trees (28) may be employed. These are particularly useful for handling missing data (where surrogate variables can be employed for node splitting) and for identification of interactions (for example between treatment and age).

Mechanistic Evaluation

As indicated in the protocol (24), a mechanistic evaluation will also be undertaken, comprising direct investigations into the relationships between the assay quantities for the various DNA fragmentation and chromatin compaction assays (including HBS) and the clinical outcomes. These will include *t*-tests, chi^2 tests and other tests of association and correlation as required. As well as the assay values themselves, derived variables, specifically DNA fragmentation and Chromatic compaction, as shown in Figures 1 and 2, will be considered. Since these latent variables are defined by multiple measures, they should have relatively lower measurement error.

Conclusion

With this SAP we present the analyses that will be published in the primary publication for the clinical aspect of the trial. By publishing this SAP before prior to unblinding of any investigators, we avoid any bias that may arise from knowledge of outcome and data-driven results.

The aim of the HABSelect study is to compare the use of PICSI to PVP-ISCI procedures for treatment of male fertility. With the publication of this paper pre-specifying the analyses to be used, we hope that the results from the HABSelect trial will be as transparent as possible.

Abbreviations

IVF: In Vitro Fertilisation; ART: Assisted Reproduction Technologies; ICSI: Intra-Cytoplasmic Sperm Injection; DGC: Density Gradient Configuration; PVP: Polyvinylpyrrolidone; PVP-ISCI: Polyvinylpyrrolidone Intra-Cytoplasmic Sperm Injection; PISCI: PISCI (hyaluronic acid coated plated) Selected Intra-Cytoplasmic Sperm Injection; HABSelect: Hyaluronic Acid Binding Sperm Selection; NICE: National Institute of Clinical Excellence; SAP: Statistical Analysis Plan; HBS:

Hyaluronic Acid Binding Score; DMEC: Data Monitoring and Ethics Committee; HFEA: Human Fertilisation and Embryology Authority; ITT: Intention To Treat; MAR: Missing At Random; SD: Standard Deviation; IQR: Inter-quartile Range; FSH: Follicle Stimulating Hormone; AMH: Anti-Mullerian Hormone

References

1. Meseguer, M. *et al.* Effect of sperm DNA fragmentation on pregnancy outcome depends on oocyte quality. *Fertil Steril* **95**, 124-128.

2. WHO. (ed. W.H. Organization)2010).

3. Zini, A., Bielecki, R., Phang, D. & Zenzes, M.T. Correlations between two markers of sperm DNA integrity, DNA denaturation and DNA fragmentation, in fertile and infertile men. *Fertil Steril* **75**, 674-677 (2001).

4. Simon, L., Lutton, D., McManus, J. & Lewis, S.E. Sperm DNA damage measured by the alkaline Comet assay as an independent predictor of male infertility and in vitro fertilization success. *Fertil Steril* **95**, 652-657 (2010).

5. Sousa, A.P. *et al.* Dual use of Diff-Quik-like stains for the simultaneous evaluation of human sperm morphology and chromatin status. *Hum Reprod* **24**, 28-36 (2009).

6. Micinski, P., Pawlicki, K., Wielgus, E., Bochenek, M. & Tworkowska, I. The sperm chromatin structure assay (SCSA) as prognostic factor in IVF/ICSI program. *Reprod Biol* **9**, 65-70 (2009).

7. Gu, L.J., Chen, Z.W., Chen, Z.J., Xu, J.F. & Li, M. Sperm chromatin anomalies have an adverse effect on the outcome of conventional in vitro fertilization: a study with strictly controlled external factors. *Fertil Steril* **92**, 1344-1346 (2009).

8. Zini, A., Boman, J.M., Belzile, E. & Ciampi, A. Sperm DNA damage is associated with an increased risk of pregnancy loss after IVF and ICSI: systematic review and meta-analysis. *Hum Reprod* **23**, 2663-2668 (2008).

9. Lewis, S.E., Agbaje, I. & Alvarez, J. Sperm DNA tests as useful adjuncts to semen analysis. *Syst Biol Reprod Med* **54**, 111-125 (2008).

10. Bhattacharya, S.M. Association of various sperm parameters with unexplained repeated early pregnancy loss--which is most important? *Int Urol Nephrol* **40**, 391-395 (2008).

11. Bakos, H.W., Thompson, J.G., Feil, D. & Lane, M. Sperm DNA damage is associated with assisted reproductive technology pregnancy. *Int J Androl* **31**, 518-526 (2008).

12. Zini, A. & Libman, J. Sperm DNA damage: clinical significance in the era of assisted reproduction. *CMAJ* **175**, 495-500 (2006).

13. Lewis, S.E. & Aitken, R.J. DNA damage to spermatozoa impacts on fertilization and pregnancy. Cell Tissue Res 322, 33-41 (2005).

14. Huszar, G. *et al.* Cytoplasmic extrusion and the switch from creatine kinase B to M isoform are completed by the commencement of epididymal transport in human and stallion spermatozoa. *J Androl* **19**, 11-20. (1998).

15. Moretti, E. *et al.* Relationship among head size, morphology and chromosome structure in human spermatozoa. *Fertility and Sterility*, 138 (1997).

16. Parmegiani, L. *et al.* Efficiency of hyaluronic acid (HA) sperm selection. *J Assist Reprod Genet* **27**, 13-16 (2010).

17. Worrilow KC, Eid S, Woodhouse D, Perloe M, Smith S, Witmyer J, et al. Use of hyaluronan in the selection of sperm for intracytoplasmic sperm injection (ICSI): significant improvement in clinical outcomes--multicenter, double-blinded and randomized controlled trial. *Hum Reprod*. Nov 30.

18. National Institute of Clinical Excellence. Review of Fertlity Guidance, http://www.nice.org.uk/newsroom/pressreleases/NICEOutlinesReviewOfFertilityGuideline.jsp

19. Prinosilova, P. *et al.* Selectivity of hyaluronic acid binding for spermatozoa with normal Tygerberg strict morphology. *Reprod Biomed Online* **18**, 177-183 (2009).

20. European Society of Human Reproduction and Embryology. Art Fact Sheet, http://www.eshre.eu/sitecore/content/Home/Guidelines%20and%20Legal/ART%20fact%20sheet.

21. White IR, Thompson SG Adjusting for partially missing baseline measurements in randomised trials. Stat Med 2005; **24**:993-1007

22. White, IR. Sensitivity analysis for randomised trials with missing outcome data. UK Stata Users' Group 2011. Retrieved from http://www.stata.com/meeting/uk11/abstracts/UK11_White.pdf

23. British Medical Journal. Multiple Imputation for Missing Data in Epidemiological and Clinical Research: Potential and Pitfalls. BMJ 2009;338:b2393

24. Miller, D. HABSelect: Selection of sperm for Assisted Reproductive Treatment by prior hyaluronic acid binding: increasing live birth outcomes and reducing miscarriage rates – multicentre randomised controlled, blinded trial. Version 3.0. 2014.

25. Peduzzi, P *et al.* A simulation study of the number of events per variable in logistic regression analysis. *J. Clin. Epidemiology,* 49, 1373–1379. (1996).

26. Nelson SM, Lawlor DA. Predicting live birth, preterm delivery, and low birth weight in infants born from in vitro fertilisation: a prospective study of 144,018 treatment cycles. *PLoS medicine* 8(1):e1000386. (2011).

27. Rosseel Y. lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software, 48(2), 1–36. URL http://www.jstatsoft.org/v48/i02/. (2012).

28. Hothorn T, Hornik K, and Zeileis A. Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*, 15(3), 651–674. (2006).

29. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. (2016).

HAB Select

Report on statistical analysis of mechanistic study

Version 0.3

Date: 10 Dec 2017

Table of Contents

Introduction

This document reports the results of the statistical analysis of the HAB Select mechanistic study. The methods and approach for this were outlined in the Statistical Analysis Plan (SAP) for the Mechanistic Study.

This is an observational study without the benefit of randomization although the mechanistic cohort was sampled from couples participating within a randomized controlled trial (HAB Select).

The results are reported here following the underlying biological pathway of the treatment provided to couples. All analysis was undertaken by Robert West, University of Leeds.

Changes from the planned analysis

The plan for analysis outlined in the SAP was intended as a guide only. It gave examples of the type of analysis envisaged, the exact analysis could not be specified. A wide range of analyses could evolve following initial findings.

The SAP suggested that structural equation modelling would form the backbone of the analysis but with fewer assays undertaken, the role of structural equation modelling was much reduced. It remains a key step along the logical analysis path described here within this report but key results are better identified with other methods: classification trees and generalized regression models.

Mechanistic cohort

The mechanistic cohort of 1247 couples was sampled from the HAB Select trial of 2766 couples undergoing fertility treatment. For those couples within the mechanistic cohort, assays were performed on aliquots of ejaculate. Due to limitations of time and cost, not all couples could have ejaculate processed. In addition, there was not always sufficient ejaculate remaining from fertility treatment to enable any, or all, of the assays to be performed.

Sampling

Stratified sampling was used to select couples from the trial for participation in the mechanistic study. Stratification was by known outcome of miscarriage. This ensured that the study was enriched for miscarriage which became the primary focus once the trial had ascertained that there was no statistically significant association between treatment allocation and the primary outcome of full-term live birth.

Assays undertaken

The number of aliquots processed for each assay are given in Table 1.

Table 1 Assay counts

Number of samples	Comet	Tunel	AO	AB	Halo area	Halo ratio	Number of missing assays
131	1	1	1	1	1	1	0
76	1	1	0	1	1	1	1
12	1	0	1	1	1	1	1
74	1	1	1	0	1	1	1
137	1	1	1	1	0	0	2
9	1	0	0	1	1	1	2
62	1	1	0	0	1	1	2
26	0	1	1	0	1	1	2
10	1	0	1	0	1	1	2
109	1	1	0	1	0	0	3
3	0	1	1	1	0	0	3
22	1	0	1	1	0	0	3
16	0	1	0	0	1	1	3
113	1	1	1	0	0	0	3
9	1	0	0	0	1	1	3
3	0	0	1	0	1	1	3
14	0	1	0	1	0	0	4
23	1	0	0	1	0	0	4
1	0	0	1	1	0	0	4
43	1	1	0	0	0	0	4
22	0	1	1	0	0	0	4
3	0	0	0	0	1	1	4
29	1	0	1	0	0	0	4

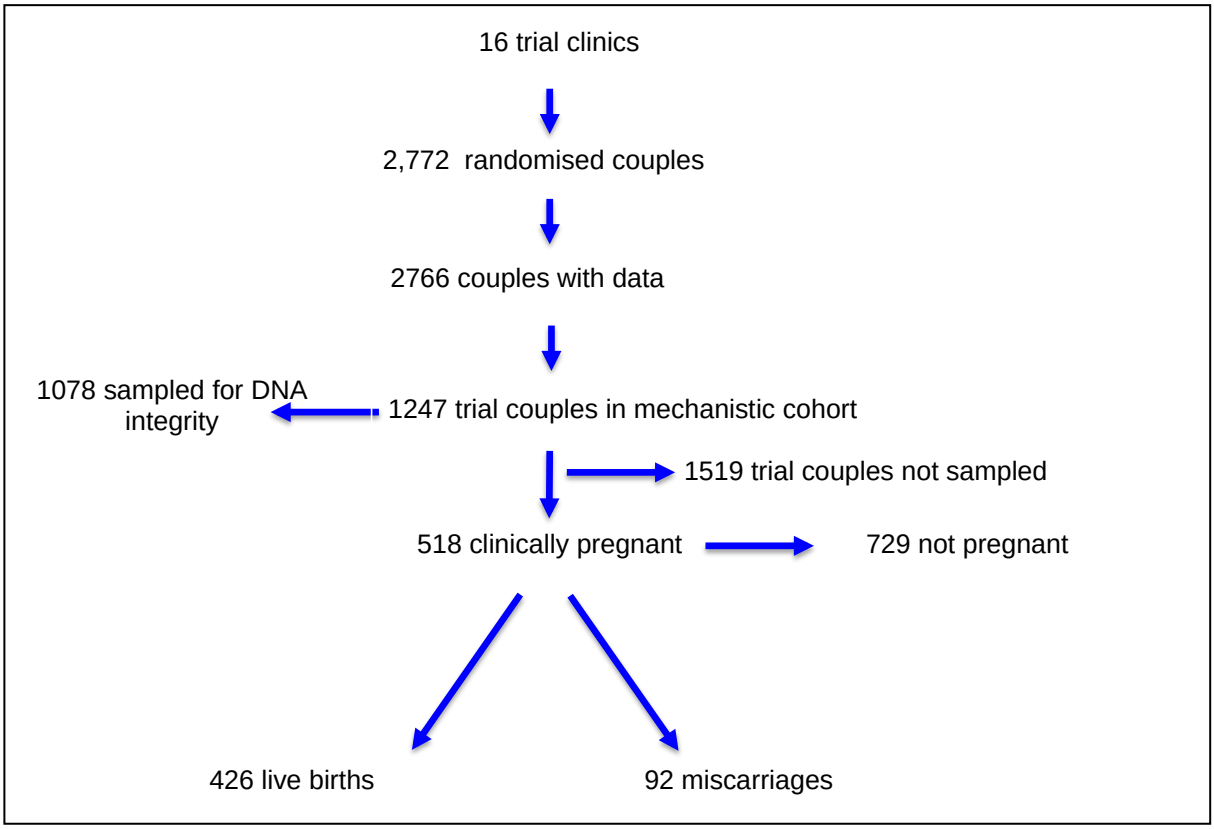
12	0	0	0	1	0	0	5
63	0	1	0	0	0	0	5
46	1	0	0	0	0	0	5
10	0	0	1	0	0	0	5
169	0	0	0	0	0	0	6
Totals	905	889	593	549	431	431	3798
Missing	342	358	654	698	816	816	3684

Table 1 shows that there are 131 couples for whom all six assay types were undertaken although 5 did not have HBS (required to complete the SEM). There were 1247-342=905 Comet assays, 1247-358=889 Tunel assays, 1247-654=593 AO assays, 1247-698=549 AB assays, 1247-816=431 Halo area and 431 Halo. Note that there are 169 couples selected for the mechanistic for whom no assay results were recorded.

Clinically pregnant subset

For some research questions, the whole mechanistic cohort is relevant, but as the biological process is followed, the analysis focusses on only those who are clinically pregnant. Hence a clinically pregnant subset is relevant. Figure 1 shows the flow of couples and the cohort or subset relevant to each analysis.

Figure 3 Flow chart for cohort subsets



Characteristics of cohort and clinically pregnant subset

The characteristics of the mechanistic cohort are tabulated in Table 2, stratified by clinical pregnancy.

Table 2 The mechanistic cohort

Characteristic	Not pregnant	Clinically pregnant	P-value
Number	729	518	
Male partner			
Age, years (mean, sd)	36.42 (5.62)	35.59 (5.39)	0.010
White (n, %)	581 (79.7%)	413 (79.7%)	0.999
Paternal BMI (mean,sd)	26.80 (4.33)	27.42 (4.38)	0.214
Alcohol units (mean, sd)	7.78 (6.18)	8.06 (7.39)	0.600
Sperm count (mean, sd)	26.91 (35.74)	25.78 (34.24)	0.579
Sperm volume (mean, sd)	2.95 (1.48)	2.98 (1.54)	0.674
Motility (mean, sd)	41.43 (19.55)	41.48 (19.45)	0.966
Female			
Female age (mean, sd)	34.06 (4.36)	33.39 (4.10)	0.006
White (n, %)	564 (77.4%)	407 (78.6%)	0.555
Female BMI (mean, sd)	24.33 (3.53)	24.45 (3.51)	0.551
Alcohol units (mean, sd)	5.32 (4.21)	5.00 (4.29)	0.371
Previous fertility treatment	259 (35.5%)	168 (32.4%)	0.283
Previous natural pregnancy	166 (22.8%)	119 (23.0%)	0.988
Previous miscarriage	126 (17.3%)	72 (13.9%)	0.125
FSH (mean, sd)	7.14 (2.20)	6.95 (2.09)	0.216
AMH (mean, sd)	20.62 (19.06)	23.32 (16.12)	0.057
Num blastocysts (mean, sd)	1.49 (0.50)	1.49 (0.53)	0.934
Assays			
HBA score (mean, sd)	74.47 (24.29)	73.38 (24.45)	0.476
Allocated PICSI (n,%)	365 (50.1%)	261 (50.4%)	0.958
TnI (mean, sd)	12.42 (15.60)	12.18 (13.81)	0.812

AO (mean, sd)	45.30 (15.48)	45.68 (15.73)	0.773
Comet (mean, sd)	19.02 (9.38)	18.77 (9.79)	0.698
Halo area (mean, sd)	168.40 (64.09)	179.66 (61.31)	0.067
Halo ratio (mean, sd)	3.67 (1.55)	3.76 (1.66)	0.555
AB (mean, sd)	64.76 (21.54)	61.19 (23.00)	0.066

Table 3 The clinically pregnant subset

Characteristic	Miscarriage	Live birth	P-value
Number	92	426	
Male partner			
Age, years (mean, sd)	36.54 (5.85)	35.39 (5.27)	0.063
White (n, %)	72 (78.3%)	341 (80.0%)	0.808
Paternal BMI (mean,sd)	26.77 (3.59)	27.57 (4.54)	0.271
Alcohol units (mean, sd)	8.12 (7.20)	8.04 (7.46)	0.942
Sperm count (mean, sd)	25.53 (31.87)	25.83 (34.76)	0.938
Sperm volume (mean, sd)	2.74 (1.76)	3.04 (1.49)	0.092
Motility (mean, sd)	40.62 (18.28)	41.66 (19.70)	0.655
Female			
Female age (mean, sd)	34.65 (4.20)	33.12 (4.03)	0.001
White (n, %)	68 (73.9%)	339 (79.6%)	0.289
Female BMI (mean, sd)	24.77 (3.73)	24.38 (3.46)	0.331
Alcohol units (mean, sd)	4.59 (3.32)	5.09 (4.50)	0.464
Previous fertility treatment	32 (34.8%)	136 (31.9%)	0.683
Previous natural pregnancy	20 (21.7%)	99 (23.2%)	0.862
Previous miscarriage	10 (10.9%)	62 (14.6%)	0.447
FSH (mean, sd)	7.16 (2.64)	6.91 (1.94)	0.381
FAMH (mean, sd)	22.67 (14.93)	23.46 (16.41)	0.753
Num blastocysts (mean, sd)	1.34 (0.68)	1.49 (0.53)	0.001
Assays			
HBA score (mean, sd)	75.79 (21.06)	72.89 (25.08)	0.357
Allocated PICSI (n,%)	32 (34.8%)	229 (53.8%)	0.001
Tnl (mean, sd)	10.24 (10.34)	12.57 (14.39)	0.235

AO (mean, sd)	49.60 (14.12)	44.86 (15.96)	0.076
Comet (mean, sd)	20.68 (10.11)	18.36 (9.68)	0.087
Halo area (mean, sd)	178.93 (56.22)	179.80 (62.40)	0.945
Halo ratio (mean, sd)	3.41 (1.30)	3.83 (1.72)	0.220
AB (mean, sd)	63.89 (23.62)	60.55 (22.87)	0.404

Inspection of assay values

The initial analysis starts with exploring the association between each assay and the HBA score. This was done by fitting simple linear regression models and noting the significance of the slope term. Results are tabulated in Table 4.

Table 4 Associations between HAB score and assay values

Assay	n	Estimate	SE	t-value	p-value
Tnl	810	-0.076	0.054	-6.99	<0.001
AO	555	-0.237	0.067	-3.84	<0.001
Comet	854	-0.202	0.083	-2.45	0.015
Halo area	406	0.068	0.011	6.05	<0.001
Halo ratio	406	1.953	0.455	4.29	<0.001
AB	514	0.094	0.048	1.98	0.049

As all assays measure DNA damage by one means or another, it was anticipated that assay results would be correlated (see Table 5). A multivariable regression was therefore fitted to establish the key assays for predicting HBA score. See Table 6 for the coefficients of the multivariable regression.

Table 5 Table of correlations

Assay	Tnl	AO	Comet	Halo area	Halo ratio	AB
Tunel	1					
AO	0.01	1				

Comet	0.05	0.01	1			
Halo area	-0.17	0.10	-0.03	1		
Halo ratio	-0.10	-0.03	0.17	0.54	1	
AB	0.14	0.26	0.06	0.02	-0.13	1

The correlations in Table 5 are not very strong suggesting that the assays may measure different aspects of DNA fragmentation (Tnl, AO, Comet) and compaction (AB and Halo).

Table 6 Table of coefficients for the regression of HBA score on assay values

Assay	n	Estimate	Std error	t-value	p-value
(Intercept)		95.240	5.681	16.77	<0.001
Tunel		-0.165	0.060	-2.75	0.007
AO		-0.129	0.070	-1.84	0.068
Comet		-0.160	0.091	-1.77	0.079
Halo area		0.027	0.020	1.34	0.184
AB		-0.002	0.042	-0.041	0.967
Halo ratio		0.020	0.776	0.025	0.980

It should be noted that the assay values all have an arbitrary scale and central value. Consequently, their absolute values have no meaning but, for example a higher Tunel value indicates more DNA fragmentation.

The results from Table 6 suggest that Tunel, AO and Comet, which all measure fragmentation have potential for prediction of HBA score and outcomes downstream in the biological process.

Structural equation modelling

Structural equation modelling was the primary analysis described in the SAP for the mechanistic study. It was intended that measurement models were built for fragmentation (measured by Tunel, AO and Comet) and for compaction (measured by AB and Halo). The regression modelling above

supports the work with Tunel, AO, and Comet measuring an underlying latent variable which can be named fragmentation. The modelling is graphically represented in Figure 2. The package lavaan version 0.5 was used within the statistical environment R version 3.3.2.

For the SEM, the model converges, using 126 complete observations, leaving 5 degrees of freedom. Overall chi-sq fit is 10.022 so that fit significance is p=0.075.

A table of coefficients from the SEM is given in Table 7.

Table 7 Coefficients from the SEM for fragmentation and compaction predicting HBA score

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)
frag =~				
tnl.frag	1.000			
ao.frag	0.650	0.313	2.073	0.038
comet.frag	0.370	0.215	1.723	0.085
comp =~				
ab	1.000			
halo.area	-0.951	0.908	-1.048	0.295

Regressions:

	Estimate	Std.Err	z-value	P(>|z|)
hba_score ~				
frag	-0.477	0.301	-1.586	0.113
comp	-0.633	0.259	-2.450	0.014

The results are sufficient to show that the underlying premise, that is the set of theoretical relationships illustrated within the SAP, was viable.

The first research question is:

RQn1: Can useful expressions be derived from the assays to represent the aspects of fragmentation and compaction of DNA within the sperm samples?

In more direct form, the latent variables of *fragmentation* and *compaction* might be defined as follows:

Fragmentation = 0.495***Tunel* + 0.322***AO* + 0.183***Comet*

Compaction = 0.244***Halo.area* – 0.256***AB* + 7.5

The arbitrary constant 7.5 and and some rescaling are applied in the compaction equation in order to provide a scale similar to that of fragmentation. Note now that *fragmentation* ranges from 10.2 to 66.8 and *compaction* from 1.4 to 67.4.

Note that since scales and centrality of the assay values were arbitrary, it is not possible to interpret the size of the coefficients. With only 126 complete observations, it was not anticipated that statistical significance of all coefficients would be seen. The significance values provided though do encourage further investigation with a larger dataset now that model viability has been confirmed.

With the above definitions for *fragmentation* and *compaction* the regression predicting *HAB score* becomes

HAB score = 99.2 – 0.71***fragmentation*

where the compaction term is dropped since it is not statistically significant and contributes little.

So that HAB score decreases with fragmentation.

For the remainder of the analysis the above definitions of *fragmentation* and *compaction* will be used to express two dimensions of the information from the six assays.

Motility, sperm count, and sperm concentration are predicted from the following models:

Motility = 18.8 + 0.21***HAB score* + 0.26***compaction*

Sperm count = 61.5 – 1.03**fragmentation*

Sperm concentration = 28.7 – 0.51* *fragmentation*

So that *motility* increases with *compaction* and *HAB score* and *sperm count* and *sperm concentration* decrease with *fragmentation*. Terms that were not statistically significant were dropped from the models.

Plots of the relationships are shown in Figures 3-6.

Figure 4 Plot of motility against HBA score

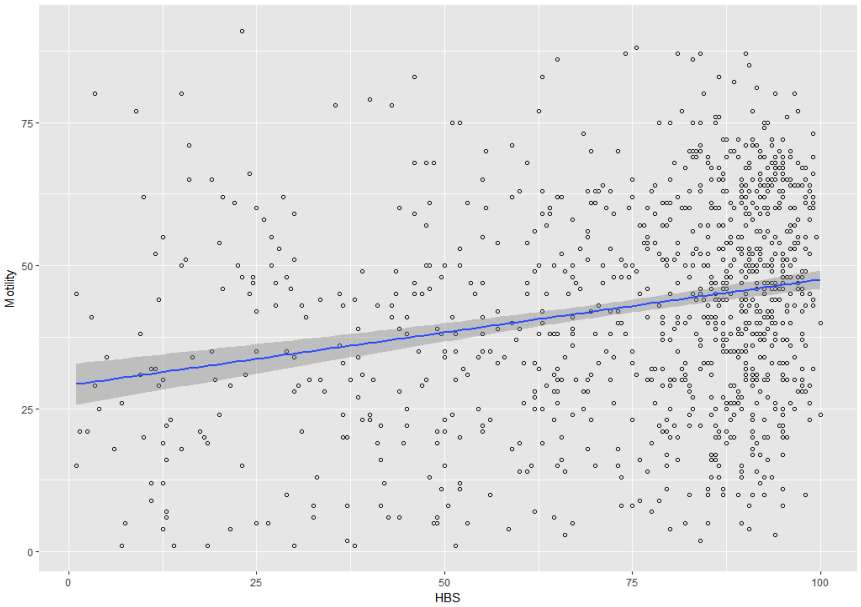


Figure 5 Plot of motility against compaction

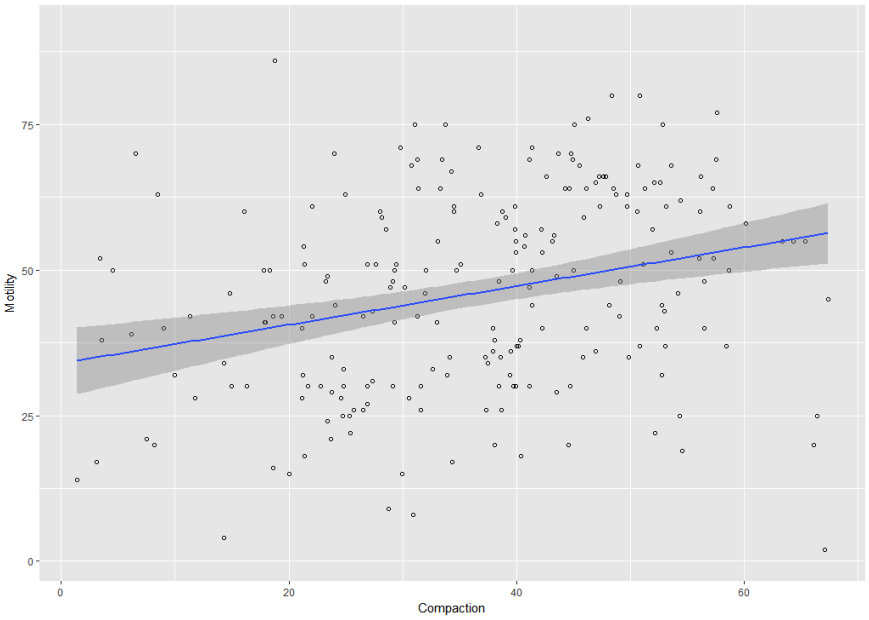


Figure 6 Plot of sperm count against fragmentation

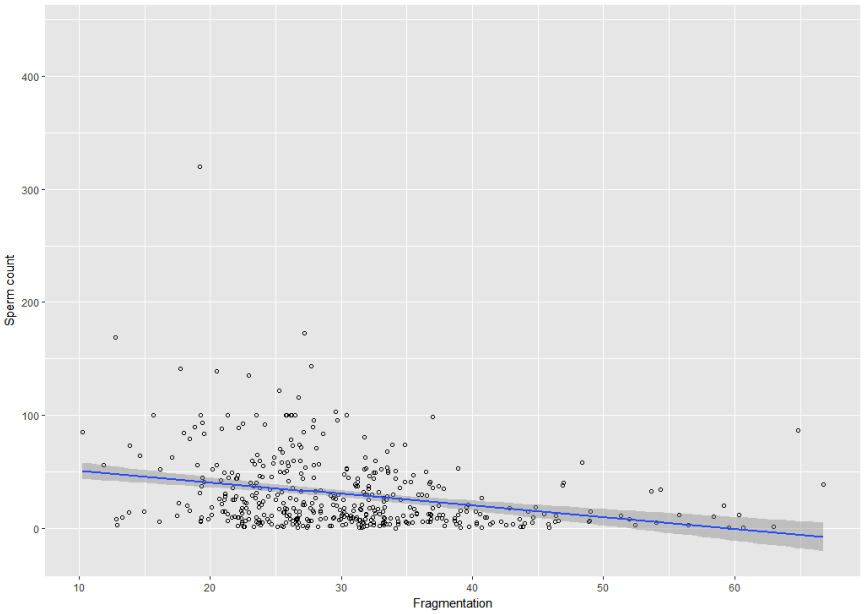
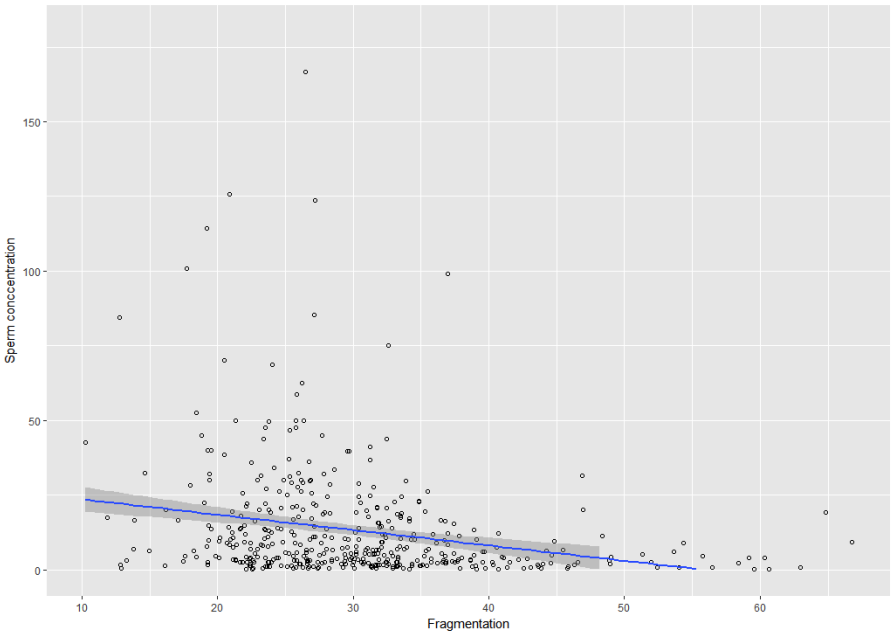


Figure 7 Plot of sperm concentration against fragmentation



In the fitted models each of the relationships in the four graphs (Figures 3-6) was assumed to be linear. A generalised additive model, with spline functions, was fitted for each relationship which permitted non-linear terms. For each situation, the extra complexity was not deemed to be sufficiently advantageous and linear fits only were used. That is nonlinear relationships were regarded as over-fitting. The predictive models were therefore taken as satisfactory.

Fertilisation

At the start of the biological process, eggs taken from the female partner are fertilised using either ICSI or PICSI. The second research question is:

RQn2: Is the fertilisation rate of eggs associated with either the allocation to treatment, the HBA score, fragmentation, or compaction?

Records were kept of the number of eggs fertilised and the successful growth of those eggs into embryos. For a couple, the number of eggs fertilised varied from 1 to 35 and the number of embryos ranged from 0 to 23. A binomial regression, with a log link function, of the fertilisation rate based on successful and failed fertilisation was undertaken to answer RQn2. The coefficients from that regression are given in Table 8.

Table 8 Coefficients from the binomial regression of fertilisation rate

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.3270	0.0439	-7.455	<0.001	***
PICSI	-0.0377	0.0105	-3.601	<0.001	***
HBS_score	0.0010	0.0002	4.775	<0.001	***
Female age	-0.0029	0.0012	-2.346	0.019	*

Table 8 gives from 1888 couples. Terms for fragmentation and for compaction were dropped as far from statistically significant. Since there was no dependence than on assays, the full HAB Select trial dataset could be used rather than restricting to the mechanistic cohort. Conversion of the coefficients to Relative Risk (RR) is shown in Table 9.

Table 9 Relative risk of treatment, HBA score and age of female partner for fertilisation of eggs with 95% confidence intervals

	RR	(Lower, Upper)
PICSI	0.9630	(0.9434, 0.9829)
HBA_score	1.0010	(1.0006, 1.0015)
Female age	0.9971	(0.9947, 0.9995)

An increased HBA score is associated with a higher fertilisation rate, PICSI reduces the fertilisation rate and the fertilisation rate decreases with female age.

Clinical pregnancy

After fertilisation and implantation, the next stage of the biological process of interest to study is clinical pregnancy, the research question being:

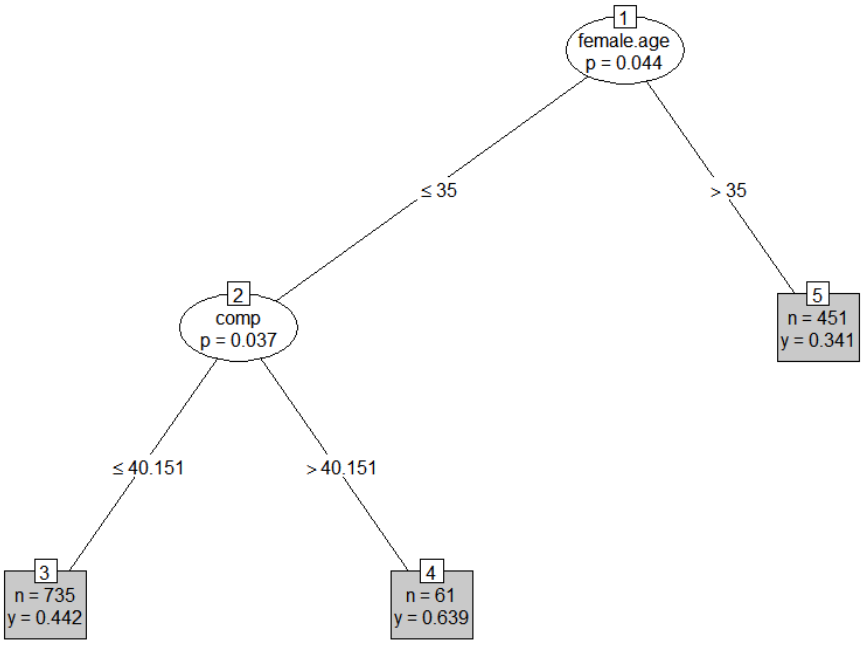
RQn3: Which factors are associated with clinical pregnancy? In particular, are HAB score, fragmentation or compaction associated with changes in pregnancy rates?

The question was answered through two different statistical methods: by classification tree and by logistic regression. The classification tree had the advantage of being able to handle highly correlated covariates such as maternal age and paternal age. On the other hand logistic regression

is more familiar to researchers. By taking two approaches, a more thorough exploration of the relationship between factors and clinical pregnancy was provided.

The results from fitting a classification tree, where the outcome is clinical pregnancy yes or no are shown in Figure 7. The elliptical nodes represent branching by the variable within and the square nodes are terminal nodes where couples have been classified by chance of clinical pregnancy.

Figure 8 Classification tree for clinical pregnancy based on the mechanistic cohort. y is the proportion of clinical pregnancies at each terminal node.



Note that the classification tree predicts that the clinical pregnancy rate is 34.1% for women over 35 years but for younger women pregnancy rate depends upon compaction with the predicted rate being 63.9% for couples with high compaction (>40.151 arbitrary units) and reducing to 44.2% for couples with low compaction. Note that for those couples for which compaction is not determined, surrogate variables are used in its place. This enables use of all 1247 couples data rather than just the 228 for which compaction is measured, although confidence in this result is limited due to the high number of surogate splits that are performed.

Fitting a logistic regression model instead of a classification tree, the fitted model coefficients are given in Table 10.

Table 10 Coefficients for logistic regression of clinical pregnancy within the mechanistic cohort

```
Coefficients:
                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                1.965     1.288     1.526    0.127
Female white               1.454     0.439     3.313    0.001 ***
Compaction                 0.023     0.010     2.297    0.022 *
Female age                -0.128     0.038    -3.396    0.001 ***
```

Note that ethnicity of the female partner is added as a putative factor with women known to be white, rather than unknown or other ethnicity, being more likely to become pregnant. As with the tree pregnancy rate decreases with *age* and increases with *compaction*.

Miscarriage

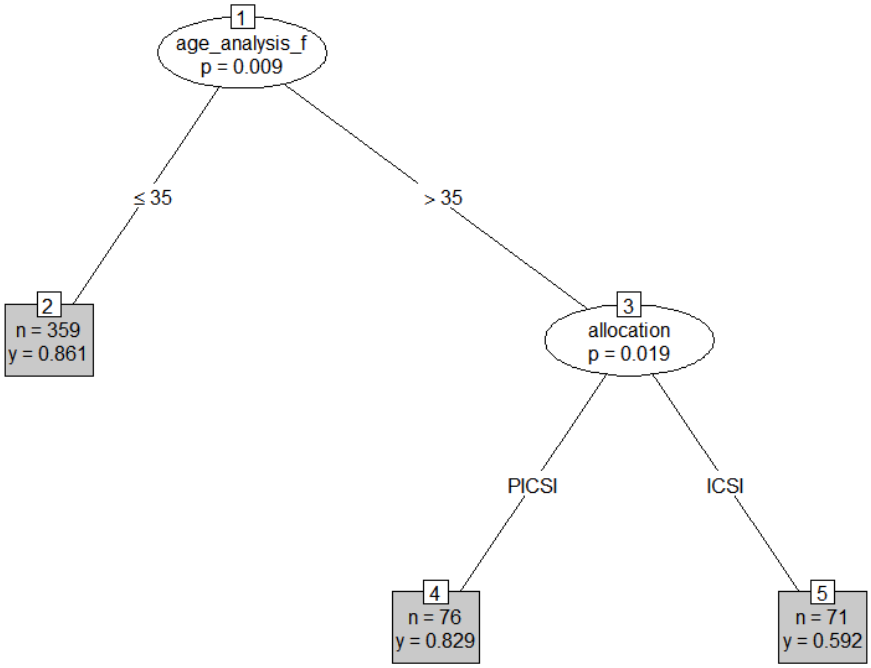
The HAB Select trial ascertained that there was insufficient evidence for an increase in live birth due to allocation to ICSI or PICSI. The focus of this mechanistic study therefore turned to miscarriage.

It should be noted that multiple blastocysts are transferred. As a consequence, one of more of these can 'miscarry' and these are recorded as miscarriages within the HAB Select trial database. The focus here however is the situation when all of the transferred blastocysts are miscarried. That is, we compare the situation where there is a live birth or not following clinical pregnancy. For convenience we define, for the work undertaken within the mechanistic cohort, *a miscarriage occurs when clinical pregnancy does not result in a live birth*. Note that by live birth we mean live birth at term (more than 37 weeks) or pre-term.

As before, modelling of the outcome, miscarriage or live birth, is undertaken first using a classification tree and then by logistic regression. In both cases, the clinically pregnant subset forms the population of interest.

The fitted tree is shown in Figure 8.

Figure 8 Classification tree for live birth or miscarriage within the clinically pregnant subset of the mechanistic study. y is the proportion of live births (term + preterm) at each terminal node.



Note that the classification tree does not use all variables, although these are made available. In particular, Table 3 shows that, in terms of a univariable analysis only, there is a significant difference between the miscarriage and live birth groups according to the number of blastocysts transferred. It appears that having accounted for age of the female partner and the allocation treatment (ICSI or PICSI), the number of blastocysts no longer discriminates.

The tree indicates that for the oldest female partners, the rate of miscarriage is 31.0%, that is for those over 37 years (1-0.69)*100 = 31%, specifically 48 of the 155 clinical pregnancies end with miscarriage and 107 of 155 with live birth.

For couples where the age of the female partner is no greater than 37 years and PICSI treatment is given, 42 of 409 miscarry, a rate of 10.3%. If for these couples ICSI is given, the rate is 29 of 108 (26.9%) for couples with female age 35-37 and 37 of 300 (12.3%) for those females age no more than 34 years. The conclusion is that PICSI is the preferred treatment for those couples where the female is aged 35-37 years.

Although 35-37 years appears a restrictive range this accounts for 23.0% of couples recruited to the HAB Select trial, and 34.3% of those that achieved clinical pregnancy.

Fitting a logistic regression achieves similar, although not identical results. There is a significant interaction between age of the female partner and the treatment allocated. Table 11 gives the table of coefficients for the final model where non-significant terms have been dropped.

Table 11 Coefficients in the logistic regression of live birth/miscarriage among those couples attaining clinical pregnancy within the mechanistic study

Coefficients:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	7.089	1.423	4.981	<0.001	***
PICSI	-5.514	2.092	-2.636	0.008	**
Female age	-0.173	0.041	-4.251	<0.001	***
PICSI:female age	0.185	0.061	3.019	0.003	**

With the interaction term it is difficult to interpret the coefficients of the logistic regression specified by Table 11. Table 12 therefore gives some examples for couples where the female partner is either 30 or 37 years and the treatment is either ICSI or PICSI.

Table 12 Table of percentages of miscarriage for the model specified in Table 11

Allocation	Female age 30	Feamle age 37
ICSI	13.0%	33.5%
PICSI	12.7%	11.8%

From Table 12 the advantage of PICSI for older women, once clinical pregnancy has been achieved is clear.

The fitted models within the mechanistic cohort for miscarriage dropped all terms relating to the assays. As a consequence, it was possible to use the full trial dataset to provide greater power for the models by using more couples.

Characteristics of the 2766 HAB Select trial participants are provided in the trial report and are not replicated here. The characteristics of couples achieving clinical pregnancy are given in Table 13.

Table 13 Those couples from the HAB Select trial achieving clinical pregnancy

Characteristic	Miscarriage	Live birth	P-value
Number	156	816	
Male partner			
Age, years (mean, sd)	36.85 (5.63)	35.37 (5.18)	0.001
White (n, %)	121 (77.6%)	642 (78.7%)	0.839
Paternal BMI (mean,sd)	26.62 (3.41)	27.39 (4.48)	0.195
Alcohol units (mean, sd)	7.66 (6.43)	7.93 (6.74)	0.725
Sperm count (mean, sd)	25.82 (31.71)	23.08 (32.60)	0.107
Sperm volume (mean, sd)	2.86 (1.60)	2.97 (1.53)	0.416
Motility (mean, sd)	40.22 (18.45)	39.79 (19.98)	0.816
Female			
Female age (mean, sd)	34.63 (4.39)	32.88 (4.02)	0.001
White (n, %)	115 (73.7%)	652 (79.9%)	0.104
Female BMI (mean, sd)	24.43 (3.43)	24.57 (3.47)	0.641
Alcohol units (mean, sd)	4.85 (3.27)	5.08 (4.39)	0.652
Previous fertility treatment	52 (33.3%)	215 (26.3%)	0.090
Previous natural pregnancy	34 (21.8%)	194 (23.8%)	0.666
Previous miscarriage	20 (12.8%)	111 (13.6%)	0.893
FSH (mean, sd)	7.14 (2.73)	7.02 (2.02)	0.596
FAMH (mean, sd)	25.80 (20.82)	24.49 (17.49)	0.523
Num blastocysts (mean, sd)	1.45 (0.51)	1.40 (0.53)	0.368
Assays			
HBA score (mean, sd)	78.59 (20.69)	72.58 (25.86)	0.021
Allocated PICSI (n,%)	60 (38.5%)	425 (52.1%)	0.002

A classification tree and a logistic regression were fitted as before for the clinically pregnant couples and the final models are presented in Figure 9 and Table 14 based on 972 couples.

Figure 9 Classifiation tree for trial couples achieving clinical pregnancy. Y is the proportion of clinical pregnancies at each terminal node.

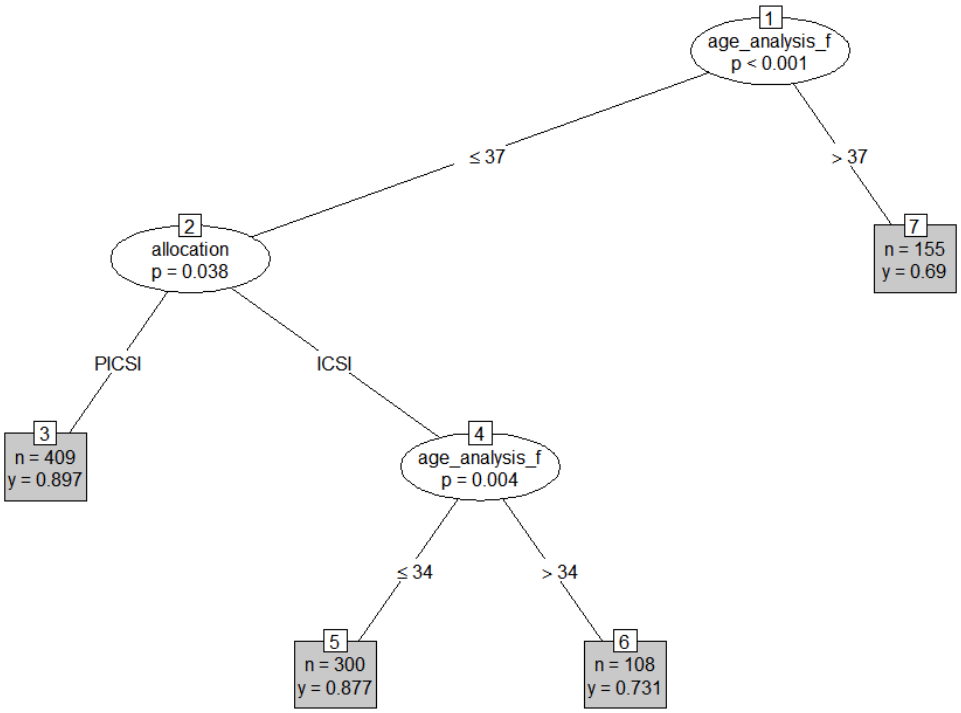


Table 14 Coefficients of logistic regression of live birth/miscarriage for those clinically pregnant couples within the HAB Select trial

Coefficients:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	7.027	1.117	6.290	<0.001	***
PICSI	-3.739	1.585	-2.359	0.018	*
Female age	-0.165	0.032	-5.164	<0.001	***
PICSI:female age	0.125	0.046	2.718	0.007	**

The classification tree and the logistic regression with the trail data reflect the findings from the mechanisic study. There is a statistically significant interaction between the allocation to treatment and the age of the female partner among those who achieve clinical pregnancy. The coefficients from Table 14 can be interpretted more easily by first considering the treatments, ICSI and PICSI. With ICSI there is a decrease in the probability of live birth with age given by the the coefficient -

0.165. When PICSI is the treatment the age term is considerably muted by adding the interaction coefficient 0.125 to the age coefficient giving a decline with age as only -0.040.

This exploratory finding might be studied further since there are implications for service delivery.

Discussion

The mechanistic study takes a different approach to the trial. The HAB Select trial is carefully designed to answer a specific research question to see if the treatment with either PICSI or ICSI as an overall service leads to more or fewer full-term live births. The mechanistic analysis on the other hand seeks to understand how the trial result comes about and explore the underlying mechanisms. For that reason, the mechanistic study breaks down the fertility service into different stages: HBA scoring, fertilisation, clinical pregnancy, and miscarriage or live birth. It is exploratory in nature and does not aim to test formal hypotheses.

HBA scoring

Six assays were considered, especially with respect to their relationships with HBA score. Prior to this study, three assays were known to be associated with DNA fragmentation within sperm: Tunel, AO, and Comet. Similarly, AB, Halo.area and Halo.ratio were linked to DNA compaction. The structural equation modelling graphically displayed in Figure 2, represents the assumed underlying assumptions, with linear relationships being indicated with the arrows in that diagram.

Table 5 showed that there were only weak correlations between assays which suggests that the assays may measure different aspects of fragmentation and compaction. Nonetheless common variation was found from the measurement models within the structural equation model and so latent variables of fragmentation and compaction could be constructed. This has the added advantage that `noise' is reduced. Each assay has measurement error, and the lack of correlation suggests that the errors (noise) are relatively large. By combining assays into *fragmentation* and *compaction*, the level of noise is reduced. Consequently subsequent analysis will have improved power due to lower noise to signal ratio although this is countered by a reduction in the number of samples that could be considered.

The modelling was a little constrained by the limited number of couples for whom all six of the assays and HAB scoring was completed (n=125), but a reasonable model fit was achieved and the assays contributed to both fragmentation and compaction as expected. It was considered appropriate therefore to continue the analysis with fragmentation and compaction.

Fertilisation

The rate of fertilisation was seen to be associated with treatment, HBA score, and age of the female partner. The rate reduces when PICSI is used rather than ICSI, decreases with female age, and increases with HBA score. Since neither fragmentation nor compaction were significantly associated a larger sample size could be used to ascertain this relationship.

This stage usually involves taking a relatively large number of eggs from the female partner so that even if the fertilisation rate is low, the overall process is not disadvantaged, there will be a sufficient number of blastocysts to transfer.

Clinical pregnancy

Both the classification tree and the logistic regression identify *age of female partner* and *compaction* as putative factors in determining the probability of clinical pregnancy. Since these are derived through two different models these can be considered as the strongest candidates. In addition, the logistic regression identifies ethnicity of the female partner as a putative factor. Note that due to the inclusion of compaction in the models, the sample size for the logistic regression was limited to n=228 couples.

Miscarriage

The HAB Select clinical trial had as primary outcome full-term live birth. The trial also regarded a miscarriage of any embryo as a miscarriage, so that for a few couples (n=8) both miscarriage and live birth were experienced. These 8 couples were excluded from the mechanistic analysis where the outcome of interest distinguished live birth (premature or full term) from miscarriage of all embryos.

The analysis for the mechanistic cohort found little evidence that either fragmentation of compaction contributed to distinguishing between live birth and miscarriage, once treatment (ICSI or PICSI) and age of female partner were known. This enabled the data from the full trial to be used in order to boost numbers studied to n=972.

Complete case analysis

There is a major limitation concerning complete case analysis. For the analyses within the mechanistic study to be valid, the assumption of missing completely at random needs to hold. Then the subset of couples for which full data is available will be representative of the whole study. Judgement is required here. The selection of which samples were sent for assay was done as a matter of convenience and might be regarded as a completely random selection. On the other hand, there were a few samples of ejaculate for which the volume was small and so the number of assays restricted. Even in these cases however the selection of assay was haphazard and might be regarded as random. There does remain the issue that some bias might have been introduced by using complete case analysis only.

No attempt was made to use imputation (that is multiple imputation) since the proportion of missing values was so high (see Table 1). When the full trial data was used for the analysis of live birth versus miscarriage, 972 of 978 clinical pregnancies were examined so that the complete case analysis would not have been influenced by the small number missing.

Linearity

For the structural equation modelling which define fragmentation and compaction as latent constructs, the linearity of the assay results was assumed. That is each of the assays was assumed to increase linearly with fragmentation and with compaction. Graphical plots revealed that this assumption was justified, and exploration with splines revealed that linear terms alone were

sufficient given the marked scatter in the relationships. Had the assays been less noisy then inclusion of nonlinear terms might have changed how fragmentation and compaction were defined and all subsequent analyses.

The regressions for fertilisation, clinical pregnancy and miscarriage/live birth also assumed linearity of the terms for age, HAB score, fragmentation, compaction. The classification trees though offer a way to check the validity of these assumptions. Since there was close agreement from two separate analysis approaches, the linearity assumptions are supported.

Conclusions

The HAB Select trial showed little evidence of differences in full-term live birth with respect to allocation of treatment (ICSI or PICSI). This has been formally tested. The trial also included a pre-defined interaction term in order to investigate differences according to age. There was little evidence to support a difference in full-term live birth by treatment by age group.

The mechanistic analysis explored different stages of the process and found putative contributions of fragmentation, compaction, HBA score and treatment (ICSI versus PICSI). These were not examined within a formal hypothesis testing framework. There could have been bias from the complete case analysis and from the assumptions of linearity, there were differences in the populations studied at each stage of the process (the trial considered only the process as a whole), and different outcomes were used. Nonetheless using two very different analysis approaches it was suggested that miscarriage rates in older women might be significantly reduced if PICSI was the treatment. This finding need to be formally explored in a further trial before sufficient evidence for this can be acknowledged. Nevertheless, this would be a simple change to the fertility service to offer PICSI as the preferred treatment for couples where the female partner is older, and could obviate the distress caused through miscarriage following a fertility cycle for a large proportion of older couples.