FACULTY OF HEALTH, AARHUS UNIVERSITY

Perceived stress and in-home assessed semen quality

A cross-sectional study

Research year report

Katrine Hjuler Lund

Department of Clinical Epidemiology, Aarhus University Hospital, Denmark

2019

Supervisors and collaborators

Ellen M. Mikkelsen, MPH, PhD, associate professor (main supervisor)
Department of Clinical Epidemiology, Aarhus University Hospital, Denmark

Anne Sofie D. Laursen, MSc, PhD, Postdoc (co-supervisor)
Department of Clinical Epidemiology, Aarhus University Hospital, Denmark

Gunnar Toft, PhD, DMSc, associate professor (co-supervisor)
Department of Clinical Epidemiology, Aarhus University Hospital, Denmark

Lauren A. Wise, professor (collaborator)
Department of Epidemiology, Boston University School of Public Health, MA, USA

Elizabeth E. Hatch, professor (collaborator)
Department of Epidemiology, Boston University School of Public Health, MA, USA

Kenneth J. Rothman, professor (collaborator)
Department of Epidemiology, Boston University School of Public Health, MA, USA
RTI, Health Solutions, Research Triangle Park, Durham, NC, USA

Tanran R. Wang, data analyst (collaborator)
Department of Epidemiology, Boston University School of Public Health, MA, USA

Greg Sommer, PhD, co-founder of Sandstone Diagnostics (collaborator)
Sandstone Diagnostics, Inc., Livermore, CA, USA

Preface

This research year report is based on a cross-sectional study of perceived stress and semen quality and a pilot study of in-home semen testing conducted during my research year in 2019 at the Department of Clinical Epidemiology (DCE), Aarhus University Hospital. My research year at DCE has been an educative and inspiring journey far beyond what I could have imagined, and I am deeply thankful for the opportunities I have been given.

A special thanks to my main supervisor Ellen M. Mikkelsen, who since our first meeting in the spring of 2018 has been very encouraging and supportive in the development and implementation of the pilot study in SnartForaeldre.dk. Throughout the research year, Ellen has opened my eyes to the field of epidemiology, and she has always provided excellent supervision – thank you for addressing all my questions and difficulties and asking challenging questions to make me reach beyond.

I am deeply grateful to Anne Sofie D. Laursen, who at all times offered me encouraging and highly valuable supervision especially with my countless questions in SAS. Further, I would like to express my thanks to Gunnar Toft for sharing extensive knowledge on male reproduction and epidemiology.

During my research year I was given the opportunity to visit our collaborators for two months at Department of Epidemiology, Boston University School of Public Health. I am deeply thankful for the hospitality and tremendous guidance from our collaborators Lauren A. Wise, Kenneth J. Rothman, Elizabeth E. Hatch, Tanran R. Wang and the rest of the PRESTO team. Further, I must express my thanks to Greg Sommer from Sandstone Diagnostics, who completed recalibration on data from both cohorts.

Finally, a special thanks to the SnartForaeldre team for interesting discussions and a great and inspiring working environment, and in particular Tina Christensen for our partnership in preparing and recruiting for the pilot study.

Katrine Hjuler Lund, March 2020

Grants

The research year was supported by the Department of Clinical Epidemiology, Aarhus University Hospital, as well as a grant from Helga og Peter Kornings Fond.

Abbreviations

BMI	Body Mass Index
CI	Confidence Interval
DAG	Directed acyclic graph
FDA	U.S Food and Drug Administration
FFQ	Food frequency questionnaire
NO	Nitric oxide
NVK	National Committee on Health Research Ethics
IQR	Interquartile range
OR	Odds ratio
PRESTO	Pregnancy Study Online
PSS	Perceived Stress Scale
SF	SnartForaeldre.dk
SF/Saedkvalitet	SnartForaeldre.dk/Saedkvalitet
Trak®	The Trak® Male Fertility System
WHO	World Health Organization

TABLE OF CONTENTS

ABSTRACT **6**

DANSK RESUMÉ **7**

MANUSCRIPT **8**

INTRODUCTION **8**

MATERIAL AND METHODS **9**

- **DESIGN AND STUDY POPULATION** 9
- **ASSESSMENT OF PERCEIVED STRESS** 9
- **ASSESSMENT OF SEMEN QUALITY** 10
- **ASSESSMENT OF COVARIATES** 10

DATA ANALYSIS **11**

RESULTS **13**

DISCUSSION **14**

ETHICAL APPROVAL **17**

TABLES **18**

FIGURES **20**

SUPPLEMENTARY **23**

PILOT STUDY OF IN-HOME SEMEN TESTING **23**

- **SNARTFORAELDRE.DK/SAEDKVALITET** 23
- **TRAK® MALE FERTILITY TESTING SYSTEM** 23
- **RECRUITMENT AND PARTICIPATION IN SF/SAEDKVALITET (AIM 1 AND 2)** 24
- **MALE PARTICIPATION (AIM 3)** 24
- **EVALUATION OF SF/SAEDKVALITET (AIM 4 AND 5)** 25

HEALTH RESEARCH ETHICS COMMITTEE **26**

MALE FERTILITY **27**

STRESS **28**

METHODOLOGICAL CONSIDERATIONS **29**

- **STUDY DESIGN** 29
- **MISSING DATA** 29

RANDOM AND SYSTEMATIC ERROR **30**

- **SELECTION BIAS** 30
- **INFORMATION BIAS** 30
- **CONFOUNDING** 31

REFERENCES **32**

SUPPLEMENTARY TABLES **35**

SUPPLEMENTARY FIGURES **37**

Abstract

Background: In North America, 18-21% of men of reproductive age report daily stress. Similarly, 23-24% of Danish men aged 16-34 years report high levels of perceived stress assessed by the 10-item Perceived Stress Scale (PSS). Some studies indicate a decline in semen quality over the past 40 years and a large study based on more than 6000 young Danish men reported a stable but high prevalence of low semen quality over the past 20 years. Results evaluating the effect of perceived stress on semen quality are inconsistent.

Aim: To examine the association between perceived stress and in-home assessed semen quality.

Methods: We used self-reported data from two ongoing prospective preconception cohort studies, Pregnancy Study Online (PRESTO) and SnartForaeldre.dk (SF). Men aged ≥21 years (PRESTO) and ≥18 years (SF) completed a baseline questionnaire on reproductive and medical history, socio-demographics, lifestyle, and stress, and were invited to perform in-home semen testing, twice with 7-10 days apart. They used the U.S Food and Drug Administration (FDA) approved test kit, Trak® Male Fertility System, which provides an in-home assessment of sperm concentration and semen volume, and a research only sperm motility test. We measured stress using the PSS (range of score 0-40) and imputed missing values at baseline and pooled data from both cohorts, N=328 (PRESTO: 276, SF:52). We performed multiple linear regression to estimate the difference within each semen parameter per 1-point higher PSS score with adjustment for potential confounders. We used the World Health Organization's (WHO) cut-off values to categorize low semen volume (<1.5ml), sperm motility (<40%), total sperm count (<39 million) and sperm concentration (<15 million/ml). We estimated the odds ratios (OR) with 95% confidence intervals (CI) of having impaired semen quality according to PSS score (<20 vs. ≥20) using logistic regression.

Results: In total, 328 men provided 576 samples. The median PSS score was 14.0 (IQR: 10.0-18.0) and 18.6% men had a PSS score ≥20. When adjusted for potential confounders the estimates (95% CI) for sperm concentration and total sperm count were 1.01 (1.00;1.03) % and 1.01 (0.99; 1.02) % higher for a 1-point higher PSS score. Semen volume was -0.01 (-0.04; 0.01) ml lower for a 1-point higher PSS score. We found no change in sperm motility (0.00 (-0.004; 0.005) %). When comparing a PSS score ≥20 vs. <20 the unadjusted ORs (95% CI) of having impaired semen volume, sperm concentration, total sperm count and sperm motility were 1.72 (0.70; 4.26), 0.91 (0.57; 1.44), 0.99 (0.60; 1.65) and 0.94 (0.67; 1.33).

Conclusion: We found a small association between perceived stress and sperm concentration and total sperm count, but no notable association for semen volume and no association for sperm motility. All analyses will be updated on a larger sample from both cohorts.

Dansk resumé

Dette tværsnitsstudie har til formål at undersøge sammenhængen mellem stress og sædkvalitet.

En stor national undersøgelse fra 2017 viste, at 23-24% af danske mænd i alderen 16-34 år ofte var meget stressede, mens en nordamerikansk rapport viste, at 19-21% af mænd i den reproduktive alder var dagligt stressede. Denne viden skal sammenholdes med diskussionen om, hvorvidt sædkvalitet er faldende. Nogle studier peger på fald i sædkvalitet, mens et dansk studie baseret på mere end 6000 unge mænd fandt en stabil sædkvalitet over de sidste 20 år, men en høj prævalens af lav sædkvalitet blandt unge mænd.

Vi anvendte selvrapporteret data fra to prospektive kohorte studier, SnartForældre.dk (SF) og Pregnancy Study Online (PRESTO). SF og PRESTO er internetbaserede studier, der rekrutterer danske og nordamerikanske par, som forsøger at opnå graviditet. Vi inkluderede mænd der havde udfyldt et baseline spørgeskema om sociodemografi, reproduktion og livsstilsfaktorer, heriblandt stress. Mændene blev inviteret til at udføre en hjemmetest af sædkvalitet ved brug af et godkendt testkit, Trak® Male Fertility System, som måler sædvolumen, koncentration og bevægelighed af sædcellerne. Mændene skulle lave to sædprøver med 7-10 dages mellemrum og analysere prøverne ved brug af Trak®. Sædkvaliteten blev beskrevet ved volumen, koncentration, total antal sædceller og sædcellernes bevægelighed.

Deltagernes stress blev målt med The Perceived Stress Scale (PSS), som er en stress skala bestående af 10 spørgsmål relateret til tanker og følelser. Ud fra skalaen beregnede vi en PSS score (skala: 0-40). Vi brugte multiple imputation til at estimere uoplyste værdier ved baseline og brugte lineær regression til at estimere forskellen i hver sædparameter for hver 1-point højere PSS score. Vi kategoriserede endvidere mændene baseret på Word Health Organizations (WHO) grænseværdier for lav sædvolumen (<1.5ml), sædcellernes bevægelighed (<40%), total antal sædceller (<39 millioner) og sædcellekoncentration (<15 millioner/ml). Vi brugte logistisk regression til at sammenligne mænd med en PSS score <20 vs. ≥20 og estimerede odds ratioen (OR) for nedsat sædkvalitet. Vi estimererede 95% konfidensintervaller (CI) og justerede for potentielle confoundere. Vi inkluderede 328 mænd (PRESTO: 276, SF: 52), hvoraf 18.6% havde en PSS score ≥20. Medianværdien for PSS-scoren var 14.0 (IQR: 10.0-18.0). Ved justering for potentielle confoundere fandt vi, at for hver 1 point højere PSS-score var sædcellekoncentrationen 1.01 (1.00;1.03) % højere og det totale antal sædceller var 1.01 (0.99; 1.02) % højere, mens sædvolumen var -0.01 (-0.04; 0.01) ml lavere. Vi fandt ingen forskel i sædcellernes bevægelighed (0.00 (-0.004; 0.005) %).

Vi fandt en lille sammenhæng mellem stress og sædcellekoncentration samt total antal sædceller, men ingen tydelig sammenhæng for sædvolumen og ingen sammenhæng mellem stress og sædcellebevægelighed. Det planlægges, at alle analyser opdateres med et større antal deltagere fra begge kohorter.

MANUSCRIPT

Introduction

The estimated prevalence of infertility is 15-20%, thus it is the most common chronic disease among individuals of reproductive age[1]. In 50% of cases, a male factor is accountable for couples' infertility[2]. Even though reported declines in sperm counts have been much debated, a meta-analysis from 2017 including 185 studies and 42,935 men found a 50-60% decline in sperm counts over the past 40 years[3]. A large Danish study based on more than 6000 young Danish men reported a stable, but high prevalence of low semen quality over the past 20 years[4]. Risk factors such as high body mass index (BMI), habitual alcohol consumption and current smoking is associated with impaired semen quality[5-7]. Further, short sleep duration in men (<6 vs. 8 hours of sleep) is related to longer time to pregnancy[8].

In 2017, the national survey "Danskernes Sundhed – Den Nationale Sundhedsprofil 2017" was published by the National Institute of Public Health, the Danish Regions and the Danish Health Authority. The national survey indicated that 23-24% of men aged 16-34 years reported high levels of perceived stress assessed by the 10-item Perceived Stress Scale (PSS)[9]. In North America, 18-21% of men of reproductive age report daily stress[10,11]. Existing literature show that stress is associated with hypertension, cardiovascular disease, obesity and depression[12-15]. In addition, a higher PSS score among women is associated with a slight increase in time to pregnancy[16].

Results from studies evaluating the effect of perceived stress on semen quality are inconsistent. One study indicated an inverse association between self-reported stress and semen quality[17], while another study among 430 Danish pregnancy planners reported no association[18]. Findings from a study assessing both perceived stress and the appearance of stressful life events suggested an inverse association between stress and semen quality[19]. Stress may affect semen quality through decreased testosterone levels, modified spermatogenesis and sexual dysfunctions[20,21]. Further, a study on medical students experiencing stress during examination suggests a potential effect via the L-arginine nitric oxide pathway[22].

To collect data on semen quality in large cohorts is expensive and cumbersome. The Trak® Male Fertility System (Trak®) is an US Food and Drug Administration (FDA) approved test kit for in-home assessment of sperm concentration and semen volume, with a research use only sperm motility test[23].

We examined the association between perceived stress and in-home assessed semen quality among men, in a subgroup of men enrolled in a Danish or North American prospective cohort of couples trying to conceive.

Material and methods

Design and study population

The study is a cross-sectional study, which uses data from two prospective cohort studies. SnartForaeldre.dk (SF) is an online ongoing prospective cohort study of Danish couples trying to conceive[24,25]. SF has recruited couples since August 2011. Pregnancy Study Online (PRESTO) is similar in design and has recruited North American couples since June 2013[26-28]. To be enrolled in SF or PRESTO men and women have to complete a screener, which confirms eligibility. Eligible women are 18-49 years (SF) or 21-45 years (PRESTO), while eligible men are ≥18 years (SF) or ≥21 years (PRESTO). Further, men and women must be in a stable relationship with a partner of the opposite sex, trying to become pregnant and not using any birth control or receiving fertility treatment.

At enrollment, men and women are invited to complete a baseline questionnaire on reproductive and medical history, socio-demographics, lifestyle and stress. Ten days after baseline, a validated food frequency questionnaire (FFQ) is available for completion. Women complete bimonthly follow up questionnaires until they report a pregnancy, stop trying to conceive or up to 12 months. Men enrolled in PRESTO are invited to pilot test in-home assessment of semen quality after baseline completion if they had tried to conceive ≤6 months with their female partner and their partner must have regular menstrual cycles. In SF, men must complete the baseline questionnaire and the FFQ before they are invited for the pilot study.

Both cohorts primarily recruit men and women through online media such as Facebook and online adds. In addition, SF recruits participants through E-box, which is a Danish communication platform between the Danish Authority and citizen.

Since the implementation of Trak® in 2015, PRESTO has invited 694 men for in-home semen testing. Among invited men, 379 (54.6%) men consented to participate, 276 (72.8%) provided one semen sample and 199 (52.5%) provided a second semen sample. From May 2019 to September 2019, SF has invited 136 men and among these, 55 (40.4%) men provided informed consent, 52 (94.5%) provided at least one semen sample and 49 (89.1%) men provided two semen samples (Figure 1).

In total, we included 328 men who provided informed consent and at least one semen sample.

Assessment of perceived stress

We assessed perceived stress by a Danish or an English version of PSS[29]. The PSS measures the extent to which individuals find their lives to be overloaded, unpredictable and uncontrollable[29]. Among respondents with at least a junior high school education, validation studies show that the PSS can capture stress experienced during the past two months[29-31]. Further, assessment of reliability via Cronbach alpha and test-retest demonstrated high internal consistency and high correlations when the test-retest was completed within a short time period. In both versions of PSS, each of the 10 items

referred to feelings and thoughts during the past month and had five response options ranging from 0 (never) to 4 (very often). We calculated a total PSS score ranging from 0–40, by summing the scores from each item. Thus, a higher score indicated a higher level of perceived stress.

Assessment of semen quality

In this study, men provided test results from two semen samples collected and analyzed at home using Trak®[32]. The men were instructed to abstain from ejaculation for 2-7 days before testing, and to collect the samples via masturbation and without the use of condoms or lubricants. In addition, they were instructed not to let semen testing interfere with their aim of achieving a pregnancy. After the in-home analysis, the men had to report the self-observed values, which were read on the prop included in the test kit, photograph and upload the test results to the study websites using a personal login and password. Finally, all de-identified photos were optically read and recalibrated by Sandstone Diagnostics, Inc., Livermore, CA.

The test kit Trak® has demonstrated adequate reproducibility and detection range for sperm concentration compared with World Health Organization (WHO) cut-off values[23], and similar evaluations are ongoing for motility and volume. In this study, we described semen quality by four parameters: semen volume (ml), sperm concentration (million/ml), total sperm count (million) and sperm motility (%).

Assessment of covariates

We obtained information on covariates from the baseline questionnaire, which included age, education, job hours per week, employment status, total household income, height and weight, alcohol consumption, smoking, sleep duration, caffeine intake, abstinence time, ever impregnated a partner, diagnosis of depression, anxiety and diabetes, fever within the past three months and intercourse frequency. We used baseline data on height and weight to calculate BMI as weight/height (kg/m^2). We examined identical covariates in both cohorts. When the test results were uploaded the participants reported abstinence time (the number of days since the most recent ejaculation) in relation to each sample.

Data analysis

We described participant characteristics by medians, interquartile range (IQR) and proportions. Given the lack of a clinical cut-off for the PSS, we categorized PSS by quartiles for the descriptive analyses (Table 1).

Semen quality (semen volume, sperm concentration, total sperm count and sperm motility) was described by the mean value of sample one and two. In case of no second sample, we used the value from the first sample. We calculated motility (%) as motile sperm concentration (million/ml)/sperm concentration (million/ml) and total sperm count (million) as sperm concentration (million/ml) x semen volume (ml).

Photos from five men in SF were unusable for calibration because of technical reasons. In these cases, we used the self-reported values or estimated lengths. Estimated lengths were provided by Sandstone Diagnostics, Inc., Livermore, CA. It is measurements of the height of the white column on the prop included in the test kit. We calculated the number of days between baseline completion and upload of the test results.

We performed multiple linear regression to estimate the difference in each semen parameter per 1-point higher PSS score. The analysis was performed in SAS using PROC GENMOD. To meet model assumptions of normally distributed residuals and linearity, we log transformed sperm concentration and total sperm count. The estimates for log-transformed variables must be interpreted as % difference for a 1-point higher PSS score. Semen volume (ml) must be interpreted as ml difference for a 1-point higher PSS score, while sperm motility (%), must be interpreted as % difference for a 1-point higher PSS score.

We identified potential confounders based on existing literature and directed acyclic graphs (DAG) (Figure 2). Thus, in our primary model (Table 2, model 1), we adjusted for cohort (SF/PRESTO) age (continuous), abstinence time (continuous), BMI (continuous), education (>15 years yes/no), current smoker (yes/no) and alcohol consumption (continuous drinks per week). Even though we did not consider abstinence time to be a confounder, we adjusted for the variable because it is strongly associated with semen quality. We considered sleep duration, diagnosis of depression and anxiety as intermediates, thus, we did not adjust for these variables in the main analyses. To evaluate the effect from sleep duration, diagnosis of depression and anxiety, we adjusted for these variables in a sub-analysis (Table 2, model 2).

We used logistic regression to estimate odds ratios (OR) with 95% confidence intervals (CI) of having impaired semen quality according to PSS score. The logistic regression was performed using PROC GENMOD. To include categories with sufficient numbers of men, we dichotomized PSS score as <20 vs. ≥20, approximately using the highest scoring quintile as cut-off[33]. We defined impaired semen

quality as being below the WHO's lower reference limit for semen volume (<1.5 ml), sperm concentration (<15 million/ml), total sperm count (<39 million) and sperm motility (<40%).

We used multiple imputation methods to account for missing data on exposure and covariates[34]. We generated five imputed datasets using PROC MI, analyzed each dataset and subsequently combined the results across the five imputed datasets using PROC MIANALYZE. Missing data on each item on the PSS ranged from 1.9-5.8% in SF, while each PSS variable in PRESTO had 0.36% missing. The majority of variables in SF had < 5% missing, while the variables "waist" and "JobHoursPerWeek" had the highest proportion of missing values (29-44%). In Presto, the majority of variables with missing values had <10% missing, but 14 variables had > 80% missing.

All analyses were performed using SAS statistical software (version 9.4, SAS Institute).

Results

Data on semen volume, sperm concentration and total sperm count were available for all men, whereas only 248 men provided data on sperm motility. In total, 80 men had missing data on sperm motility as they were enrolled in The PRESTO Semen Testing Pilot Study before Trak® was able to measure sperm motility. In total, 5 (1.5%) men had semen volume <1.5ml, 37 (11.3%) had sperm concentration <15 million/ml, 27 (8.2%) had total sperm count <39 million, and 81 (32.7%) had sperm motility <40%. The median (IQR) of semen volume was 3.8 (3.0-4.8) ml, sperm concentration had a median (IQR) of 49.0 (27.9-81.8) million/ml, while the median (IQR) of total sperm count and sperm motility were 191.1 (99.2-300.8) million and 50 (40-70) %, respectively.

The median (IQR) number of days between baseline and upload of test results from sample one was 18.0 days (10.0-31.0) and 38.0 days (28.0-59.0) between baseline and upload of test results from sample two. Assessment by cohort showed, that PRESTO had a shorter time between baseline completion and test upload. In PRESTO, the medians (IQRs) were 16.0 days (9.0-28.0) and 37.0 days (26.0-59.0), whereas the medians (IQRs) in SF were 36.0 days (28.0-43.5) and 44.0 (37.0-55.0). SF had a longer time between baseline and test upload, as men were invited for in-home semen testing after FFQ completion and not baseline completion.

The median PSS score was 14.0 (IQR: 10.0-18.0) and, 61 men (18.6%) had a PSS score ≥20.

In table 1, we present characteristics of the study population by PSS scores of <10, 11-14, 15-18 and ≥19. The distribution of age, abstinence time, alcohol intake, hours worked weekly, intercourse frequency and BMI were similar across categories of PSS. Men tended to be overweight, as the median BMI across all groups were >25. Men with a PSS score ≥19 were more likely to smoke, consume more caffeine, sleep less than 7 hours per night and have an education <15 years compared with a PSS score <10. Further, they had a higher frequency of depression and anxiety diagnoses, and the proportion of high total household income was smaller in groups with higher PSS score. When adjusted for cohort, age, BMI, abstinence time, smoking, education and alcohol intake the estimates with 95% CI for sperm concentration was 1.01 (1.00;1.03) % higher for a 1-point higher PSS score. Likewise, total sperm count was 1.01 (0.99; 1.02) % higher for a 1-point higher PSS score. Semen volume was -0.01 (-0.04; 0.01) ml lower for a 1-point higher PSS score. We found no difference in sperm motility (0.00 (-0.004; 0.005) %) (Table 2, model 1). The estimates were similar after additional adjustment for sleep duration and diagnosis of depression and anxiety (Table 2, model 2). When comparing a PSS score ≥20 vs. <20 the unadjusted ORs (95% CI) of having impaired semen volume, sperm concentration, total sperm count and sperm motility were 1.72 (0.70; 4.26), 0.91 (0.57; 1.44), 0.99 (0.60; 1.65) and 0.94 (0.67; 1.33) (Table 3).

Discussion

In this cross-sectional study, we found a small association between perceived stress and sperm concentration and total sperm count. The adjusted estimates with 95% CI for sperm concentration and total sperm count were 1.01 (1.00;1.03) % and 1.01 (0.99; 1.02) % higher for a 1-point higher PSS score. We found no notable association for semen volume, as it was -0.01 (-0.04; 0.01) ml lower for a 1-point higher PSS score. Further, there was no association for sperm motility (0.00 (-0.004; 0.005) %). To put the estimates based on our linear regression model in context, we can compare two men who are similar in all aspects but PSS score and total sperm count. If one man has a PSS score of 23 points and a total sperm count of 45 million, the other man would have a total sperm count of 45,450,000 million if having a PSS score of 24 points. It is debatable whether our findings of a 1% higher sperm concentration or total sperm count per 1-point higher PSS score is clinically relevant.

We made additional evaluations using logistic regression to estimate the odds of being classified below the WHO's cut-off values for impaired semen quality according to PSS score. We did not have a sufficient number of participants in each group for all semen parameters to permit adjustment for potential confounders. E.g. only 1.5% had a semen volume <1.5ml. The ORs were imprecise and provided inconsistent evidence for an association, although it should be noted that the dichotomization resulted in severe loss of power compared to the linear regression models. As recruitment is ongoing, all analyses will be updated on a larger sample from both cohorts to enhance precision. Further, we will evaluate the association by restricted cubic splines and an additional linear regression model treating PSS as a categorical variable.

Trak® is a convenient device to assess semen quality[32], as it does not require face-to-face appointments. While, it may be feasible and convenient, it does not account for sample loss. Men do not report on spillage, which may affect the results. A loss of the first fraction of the ejaculate has more effect on the analysis, than a loss of the last part, as the first fraction is sperm-rich[35]. However, spillage will expectedly be non-differentially misclassified, and likely bias the estimates towards the null. DNA analysis and morphology assessment are not compatible with Trak(R). However, we are still able to describe volume, concentration, total sperm count and motility.

Both cohorts enroll couples with fertility ranging from highly fertile to infertile. Unintended pregnancies are not enrolled, as they are most likely to occur among the highly fertile couples. Thus, men in our study may have a lower semen quality compared to men in the general population. Nevertheless, we assume they are unaware of their semen quality as they do not receive fertility treatment and semen analysis are mostly done to investigate infertility and for other medical reasons. As data on stress is collected prior to semen sample collection, differential misclassification of the exposure is unlikely. Further, it is unlikely that online recruitment affects our findings, unless perceived stress and semen quality differs according to internet access.

Stress can be acute and chronic. A short period of stress is often a natural and appropriate reaction to handle short term stressors e.g. a sports tournament or an exam, whereas chronic stress can have serious social, psychological or health related consequences. A review of lifestyle and male fertility suggests that permanently high levels of glucocorticoids in testes during chronic stress may induce apoptosis of sperm cells, which leads to impaired semen quality[36]. Likewise, a study on medical students by Eskiocak et al. suggests a potential link between stress and semen quality via the L-arginine-nitric oxide pathway[22]. The results reported by Eskiocak et al. indicates an inverse association between the concentration of nitric oxide and sperm concentration and percentage of rapid progressive motility of spermatozoa when comparing a stress period to a non-stress period.

However, those potential mechanisms do not explain our findings, which indicate a 1% higher sperm concentration and total sperm count per 1-point higher PSS score. Adjusting for potential confounders in the linear regression model did not change the estimates substantially. Thus, it is unlikely that the association is confounded by these variables. Nevertheless, unmeasured confounding for example by fever may affect the estimates. Men report number of fever episodes within 3 months before study entry, but we do not have any data on fever episodes in the immediate period before semen testing. We may therefore have, overestimated the association, as part of the findings could be described by fever episodes as fever affects semen quality by an increased level of reactive oxygen species, which leads to DNA damage[36].

Our findings on semen volume and sperm motility are consistent with those reported by Hjollund et al., as they found no association between male stress and semen quality[18]. However, our results suggest a 1% higher sperm concentration and total sperm count per 1-point higher PSS score. In contrast, Nordkap et al. and Janevic et al. reported inverse associations between stress and several semen parameters (Nordkap et al.: sperm concentration, total sperm count, semen volume, Janevic et al.: sperm concentration, sperm motility and morphology)[17,19]. Nordkap et al. measured stress by The Copenhagen Psychosocial Questionnaire, a four item scale[17], whereas Janevic et al. and Hjollund et al., and we used comprehensive scales based on 10 items (PSS) or 12 items (General Health Questionnaire)[18,19]. Besides using the 10 item PSS, Janevic et al. used the appearance of stressful life events, which is an objective measurement. It does not account for coping strategies and individual experiences of stress, which the PSS does. Different measurements of stress may explain the inconsistent findings across studies. Our study is comparable to Nordkap et al. and Hjollund et al., as they included Danish men (men aged 19 and pregnancy planners), whom we assume were unaware of their semen quality. Janevic et al. included men aged 38-49, that may have been aware of their semen quality (e.g. fathered a child). They may further differ from men in our study (range 28-35), as age affects male fertility[36]. Janevic et al. asked men to provide two semen samples two weeks apart and abstain from ejaculation for 2-5 days. In our study, men provided two samples 7-10 days apart with an abstinence time for 2-7 days. Differences in assessment of semen quality may also contribute to the inconsistent findings[36].

In conclusion, we found an association between perceived stress and sperm concentration and total sperm count, but the clinical relevance of a 1% difference is debatable. Our findings indicate no notable association for semen volume and no association between perceived stress and sperm motility.

Ethical approval

The Committee on Health Research Ethics in Central Denmark Region approved the SnartForaeldre.dk/Saedkvalitet study (project number 1-10-72-14-19). The Semen Testing Pilot Study was reviewed and approved by the Boston University Medical Campus Institutional Review Board (protocol number: H-31848).

Tables

Table 1. Participant characteristics of 328 men according to PSS score*

	PSS <10	PSS 11-14	PSS 15-18	PSS ≥19
No. of participants	94	83	76	75
No. of semen samples	169	149	129	129
Age, years (median, IQR)	31.5 (29.0-34.0)	31.0 (29.0-34.0)	31.0 (28.0-34.0)	32.0 (28.0-35.0)
Intercourse frequency				
<3 times /month	20.2%	30.1%	25.0%	36.0%
1 time /week	13.8%	22.9%	23.7%	20.0%
2-3 times /week	52.1%	37.3%	46.1%	30.7%
4+ times / week	13.8%	9.6%	5.3%	13.3%
Abstinence time, days (median, IQR)	3.0 (3.0-4.0)	3.0 (2.5-4.5)	3.0 (2.5-4.0)	3.0 (2.5-4.5)
BMI kg/m2 (median, IQR)	26.8 (23.7-30.9)	27.4 (23.9-31.3)	26.6 (23.7-31.0)	28.5 (24.3-32.9)
Annual household income (USD)				
<50K	4.3%	8.4%	10.5%	16.0%
50-99K	24.5%	30.1%	36.8%	50.7%
100K-149K	52.1%	38.6%	31.6%	20.0%
>150K	18.1%	21.7%	21.1%	13.3%
Refused or unknown	1.1%	1.2%	0%	0%
Education >15 years, %	76.6%	71.1%	72.4%	65.3%
Employed, %	84.0%	86.7%	80.3%	78.7%
Hours worked per week (median, IQR)	40.0 (37.0-42.0)	40.0 (40.0-45.0)	40.0 (38.3-45.0)	40.0 (40.0-50.0)
Smoking, %	10.6%	9.6%	9.2%	17.3%
Current caffeine intake, mg/day (median, IQR)	157.6 (43.4-275.0)	121.4 (44.3-282.2)	140.0 (47.1-228.0)	185.7 (80.0-295.0)
Current alcohol intake, drinks/week (median, IQR)	3.0 (1.0-7.0)	3.0 (1.0-7.0)	4.0 (0.0-9.0)	3.0 (1.0-8.0)
Sleep duration <7 hours/night, %	18.1%	33.7%	25.0%	37.3%
Ever impregnated a partner, %	33.0%	34.9%	38.2%	57.3%
Ever diagnosed with depression,%	7.4%	12.0%	6.6%	26.7%
Ever diagnosed with anxiety, %	7.4%	4.8%	2.6%	25.3%
Ever diagnosed with diabetes, %	3.2%	3.6%	2.6%	5.3%
Fever during the past 3 months, %	9.6%	12.0%	13.1%	13.3%

*Baseline characteristics are restricted to the first dataset resulting from multiple imputation.

Table 2. Difference in semen parameter per 1-point higher PSS score

			Model 1	Model 2
	N	Unadjusted (95%CI)	Adjusted (95%CI)1	Adjusted (95%CI)2
Semen volume (ml)	**328**	-0.02 (-0.04; 0.00)	-0.01 (-0.04; 0.01)	-0.01 (-0.04; 0.01)
Sperm concentration (%)3	**328**	1.02 (1.00; 1.03)	1.01 (1.00;1.03)	1.01 (0.99; 1.02)
Total sperm count (%)3	**328**	1.01 (0.99; 1.03)	1.01 (0.99; 1.02)	1.00 (0.99; 1.02)
Sperm motility (%)	**2484**	0.001 (-0.003; 0.006)	0.00 (-0.004; 0.005)	-0.00 (-0.005; 0.005)

1Adjusted for cohort, age, BMI, abstinence time, smoking, education and alcohol intake

2Adjusted for cohort, age, BMI, abstinence time, smoking, education, alcohol intake, sleep duration, diagnosis of depression and anxiety

3Estimates for sperm concentration and total sperm count must be interpreted as % difference, as they are log-transformed.

480 men in PRESTO have missing data on sperm motility

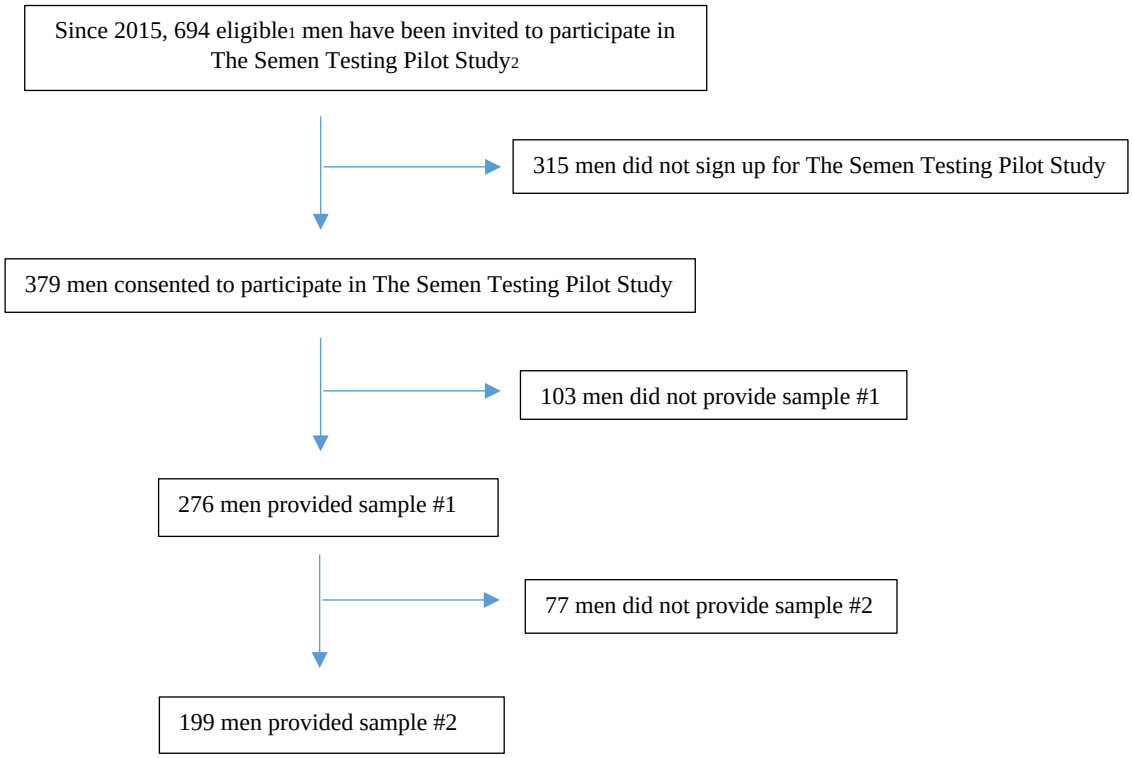
Table 3. ORs with 95%CIs of impaired semen quality according to PSS score, N=328			
	No. of men with outcome of interest	**N**	**Unadjusted OR 95%CI**
Semen volume <1.5 ml			
PSS score <20	**5**	**328**	1.0 (ref)
PSS score ≥20			1.72 (0.70-4.26)
Sperm concentration <15 million/ml			
PSS score <20	**37**	**328**	1.0 (ref)
PSS score ≥20			0.91 (0.57-1.44)
Total sperm count <39 million			
PSS score <20	**27**	**328**	1.0 (ref)
PSS score ≥20			0.99 (0.60-1.65)
Sperm motility <40%			
PSS score <20	**81**	**248**[1]	1.0 (ref)
PSS score ≥20			0.94 (0.67-1.33)

[1]80 men in PRESTO have missing data on sperm motility

Figures

Figure 1. Flow charts of enrolment

Flow chart of enrolment in The Semen Testing Pilot Study (PRESTO)



1Eligible men are: Men who completed baseline, had tried to conceive for ≤ 6 months with their female partner and their partner must have regular menstrual cycles.

2Since the beginning of The Semen Testing Pilot Study there has been times where the invitation has been suspended due to various reasons.

Flow chart of enrolment in SnartForaeldre.dk/Saedkvalitet (SF)

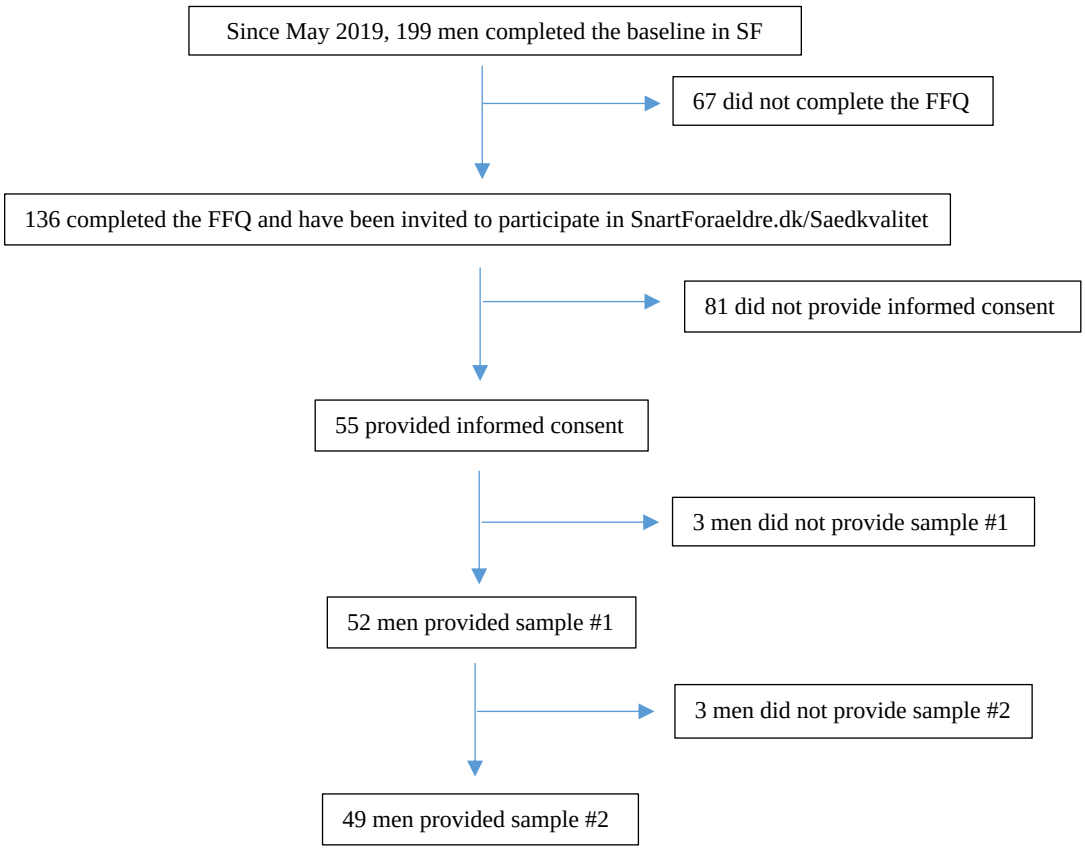
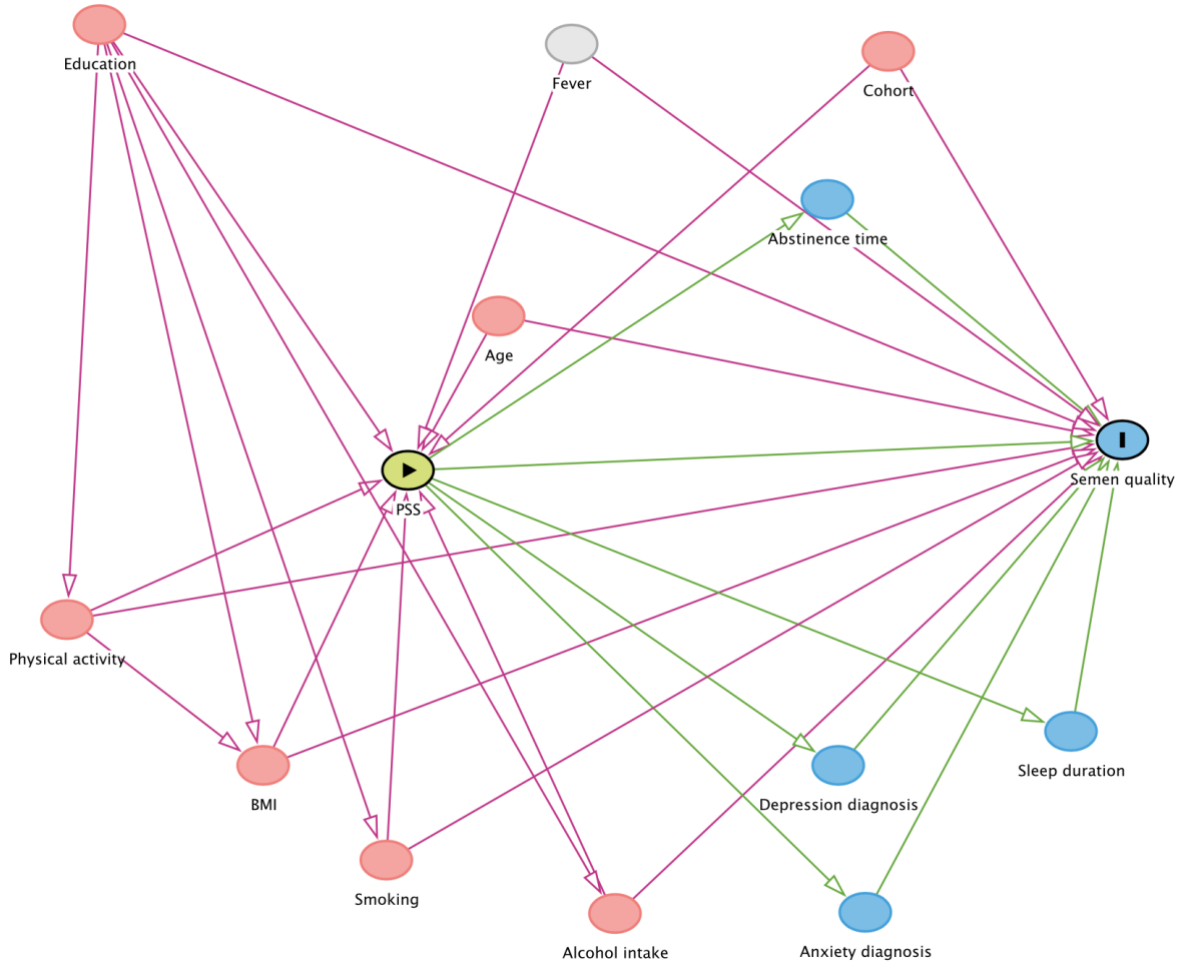


Figure 2. Directed acyclic graph



SUPPLEMENTARY

This research year rapport is based on a cross-sectional study of perceived stress and in-home assessed semen quality and the implementation of a pilot study of in-home semen testing in SnartForaeldre.dk. The Supplementary consist of a description of the test kit, recruitment and evaluation of the Danish pilot study SnartForaeldre.dk/Saedkvalitet (SF/Saedkvalitet), which was the primary aim of my research year. In addition, it contains considerations related to the cross-sectional study.

Pilot study of in-home semen testing

SnartForaeldre.dk/Saedkvalitet

In May 2019, SF launched the pilot study SF/Saedkvalitet, which aimed to evaluate the feasibility of the FDA approved device Trak® for in-home semen testing. It included the following specific aims:

1. To recruit 40 males who are willing to use the Trak® test kit
2. To assess acceptability and usability of the Trak® test kit
3. To compare male participation in SF before and after implementing the Trak® test kit
4. To compare the semen quality of Danish and American participants
5. To compare semen quality of the participants with reference values from WHO and published data on semen quality in the general population of Danish men

We asked men to collect and analyze two semen samples, 7-10 days apart using Trak®. Participants had to provide informed consent online and did not receive any financial compensation for participating. They received Trak® by ordinary mail. The participants were instructed to abstain from ejaculation for 2-7 days before testing, to collect the samples in the collection cups via masturbation and without the use of condoms or lubricants. For each sample they were asked to make two analyses using separate props to assess sperm motility and sperm count. Further, they had to report the self-observed values, which were read on the prop included in the test kit, photograph and upload the test results to the study website (www.Snartforaeldre.dk) using a personal login and password. At the study website the participants were asked to complete a usability survey. Finally, all de-identified photos were optically read and recalibrated by Sandstone Diagnostics, Inc., Livermore, CA.

Trak® Male Fertility Testing System

Trak® allows a quantitative measure of sperm concentration and semen volume, with a research use only sperm motility test[23]. The test kit has adequate reproducibility and detection range for sperm concentration compared with WHO cut-off values[23], and similar evaluations are ongoing for motility and volume.

The test kit contains two collection cups with visual marks to measure volume, a centrifuge and two cartridges (count and motility prop), disposable pipettes, sticky labels and test cards (Supplementary

figure 1 Photos of Trak®). Further, Trak® includes detailed instructions on how to collect and analyze the semen samples. After a semen sample has liquefied for 30 minutes, a small aliquot of specimen is applied to a test prop using a disposable pipette. Once the prop is sealed with a sticky label, it is placed on the centrifuge, which spins the prop for 10 minutes. For each semen sample, the participants are asked to make two analyses using separate props to assess motility and sperm count. Test results are assessed by measuring the height of the white column on the prop. Men achieve knowledge of their semen quality as they are able to read and interpret the test results by comparing the height of the white column with scales for sperm concentration and motility. The scales for interpretation of the results by the participant themselves indicate whether the semen parameter is low, moderate or high (Supplementary figure 1 Photos of Trak®).

Recruitment and participation in SF/Saedkvalitet (Aim 1 and 2)

During May 9th - September 14th, 2019 SF recruited 199 men and 981 women through E-box campaigns in three different Danish regions (North Denmark Region, Region of Southern Denmark and Region Zealand). As men mainly enroll by invitation from their female partner, the E-box campaign targeted women aged 25-35 years. Men who completed the baseline questionnaire and the FFQ were invited to participate in SF/Saedkvalitet. All participants were asked to provide informed consent online. According to the Regional Committee on Health Research Ethics in the Central Denmark Region men participating in SF/Saedkvalitet were required to receive oral information by telephone prior to providing informed consent.

Based on the two E-Box campaigns, 199 men enrolled in SF and completed the baseline questionnaire and 136 men completed the FFQ. Thus, 136 men were invited to participate in SF/Saedkvalitet. Among those, 62 (45.6%) were interested in participating and 58 (42.6%) men completed the required phone call. In total, 55 (40.4%) men provided informed consent and 52 (94.5%) men successfully uploaded results from semen sample 1 and completed a usability survey. Additionally, 49 (89.1%) men provided test results from a second sample. (Supplementary figure 2 and table 3). Thus, we succeeded to recruit more than 40 men for SF/Saedkvalitet, which was one of the aims of the pilot study.

Male participation (aim 3)

We compared male participation in SF before and after implementing the test kit by calculating male to female participation ratios using the number of men and women enrolled in each E-Box campaign. We compared ratios from E-Box campaigns in 2019 with E-Box campaigns in 2018.

In 2018, the campaigns targeted women living in the Central Denmark Region and the Capital Region of Denmark. In 2019, invitations were sent out to women living in the North Denmark Region, Region Southern Denmark and Region Zealand. We do not expect the geographical variation to have a large impact on our results, as we assume couples trying to conceive to be fairly similar in all Danish regions. In 2018, the male to female ratios were 0.21 and 0.22. Similarly, ratios in 2019 were 0.21 and 0.20. It

indicates no difference in male to female ratios across the campaigns. In conclusion, Trak® does not seem to be an incentive for men to participate in SF. Nevertheless, we consider Trak® to be usable and convenient, as 94.5% of the men who consented to participate provided at least one semen sample and positive comments on the usability survey. Thus, we will continue to recruit men for in-home semen testing in SF.

Evaluation of SF/Saedkvalitet (aim 4 and 5)

During the research year we had to change plans, as another study has reported on the feasibility of in-home semen testing[32]. This change led to a cross-sectional study of perceived stress and in-home assessed semen quality. Therefore, we did not assess all specific aims of the pilot study. We did not compare the semen quality of the participants with reference values from WHO, published data on semen quality from the general population of Danish men and American participants in The Semen Testing Pilot Study.

Health Research Ethics Committee

The health research ethics committee system is a Danish system, which consists of the National Health Research Committee (NVK) and 12 regional committees. The purpose of the committees is to ensure that studies of health research is carried out in a responsible manner. The system ensures protection of the rights, safety and well-being of subjects included in the study, while the project has the opportunity to provide new knowledge[37]. All health research projects which use personally identifiable material, e.g. tissue or cells must be notified to the committee. Clinical trials of medical equipment and products must also be notified[38].

Our pilot study SF/Saedkvalitet must be notified to a Regional Committee. Although we are not collecting, analyzing or storing any human material, the pilot study invites participants to collect and analyze semen samples at home.

We notified a Regional Committee to obtain an approval for implementing the pilot study, SF/Saedkvalitet. The application submitted for approval contained comprehensive information on how SF informs potential participants. Men invited to perform in-home semen testing received an invitation with detailed information. It underlined that participation in the pilot study is voluntary, not associated with any risks or side effects and that participation does not affect any current or future treatment in the Danish healthcare system. We provided information on the possibility of measurement errors from a defect test kit and that SF did not provide individual counselling. Likewise, men were asked to contact their general practitioner in case of any concerns regarding their fertility status. Lastly, we pointed out that participation must not affect their attempt to conceive and all participants were encouraged to read a brochure on study participation by the NVK[39].

The Regional Committee required that prior to providing informed consent, potential participants must receive oral information by SF. Thus, men with interest in the pilot study were asked to 1) call SF during specific times, 2) request a phone call from SF or 3) request a personal meeting. None of the men requested a personal meeting. The oral information was similar to what they had read online. A consent form became available online after phone call completion. Throughout the research year I have completed the majority of phone calls with potential participants prior to preparing a test kit for every participant, who signed up for the pilot study.

Male fertility

The clinical definition of infertility is the inability to conceive within 1 year of regular unprotected intercourse. Infertility is characterized by the couple and not by the individual, which makes it difficult to study[40]. In 50% of all cases a male factor is accountable for couple's infertility[2]. Birth defects, varicocele, lifestyle factors and infectious diseases are causes of male infertility. There are several ways to study male infertility. It can be assessed from studies of semen characteristics, which involves measurements of semen volume, sperm concentration, sperm motility and morphology. Further, male infertility can be accessed by studying fecundability. Fecundability is the probability of conception within one menstrual cycle.

To ensure consistency across studies analyzing semen samples, WHO have published a laboratory manual for standardized assessment of human semen[35]. According to the WHO laboratory manual, semen quality is among others affected by sample collection, activity of the accessory sex glands, size of the testis and abstinence time[35]. The duration of spermatogenesis is approximately 70 days and spermatogenesis is susceptible to effects from temperature and hormones[41]. Additionally, several risk factors such as high BMI, habitual alcohol consumption and current smoking is associated with impaired semen quality[5-7].

In 2010, WHO published new lower reference limits for semen parameters[42]. However, semen quality is highly variable, and the values should be used as guidance for evaluating a man's fertility status. Sperm concentration and total sperm count are two separate semen parameters. Sperm concentration is characterized by the number of spermatozoa per unit semen volume, while total sperm count refers to the total number of spermatozoa in the entire ejaculate. Men, whose semen parameters (total number of spermatozoa, percentages of progressive motile and morphologically normal spermatozoa) are equal to or above WHO's lower reference limit, have semen quality characterized as normozoospermia. Semen quality is characterized as azoospermia if there are no spermatozoa in the ejaculate.

In this study, we obtained data on semen quality from Trak®, which provides data on semen volume, sperm concentration and sperm motility. Another approach to assess semen quality and thereby male fertility would be analyzing sperm DNA and hormone levels.

Stress

Stress is defined in three ways in the international literature: 1) as factors in the surrounding environment, which affects the individual, 2) as an individual condition and 3) as interactions between the surrounding environment and the individual[43]. Stress can be divided into short (acute) and long-term stress (chronic). A short period of stress is often a natural and appropriate reaction to handle short term stressors e.g. a sports tournament or an exam. Long-term stress is a condition, which can have serious social, psychological or health related consequences. Each individual experience stressors differently due to differences in resources and coping strategies.

A study by Eskiocak et al.[22] based on medical students experiencing stress during exams suggests a potential mechanism between stress and semen quality via the L-arginine-nitric oxide pathway. Nitric oxide (NO) is a free radical, which is highly reactive. NO is synthesized from L-arginine via enzymes called NO synthases. However, arginine is also a substrate for the arginase enzyme, which catalyzes the transformation of L-arginine to urea and ornithine. The reaction limits the amount of L-arginine available for NO synthesis. NO is a necessary molecule for maintaining a normal sperm production and motility, nevertheless NO is also cytotoxic for cells. Eskiocak et al. found an inverse association between the NO concentration and sperm concentration, percentage of rapid progressive motility of spermatozoa and arginase activity in seminal plasma when comparing a stress period to a non-stress period. Another potential mechanism described by Ilacqua et al., suggests that permanently high levels of glucocorticoids in testes during chronic stress may induce apoptosis of cells through DNA damage, which leads to impaired semen quality[36].

Methodological considerations

Study design

In order to evaluate the association between perceived stress and semen quality, we used data from two prospective cohort studies. Cohort studies can be described as a group of people, who are classified by exposure differences. Both groups (exposed and unexposed) are followed over time to evaluate the incidence of the outcome of interest[44]. A cohort study can be conducted as a prospective or a retrospective cohort study. Prospective cohorts are assembled at present and followed into the future. PRESTO and SF are prospective cohorts, as they enroll couples trying to conceive and follow women until they report a pregnancy, stop trying to conceive or up to 12 months, whichever comes first. In retrospective cohort studies, the cohort are assembled by past records and then followed up to the present.

The current study is a cross-sectional study, which is another type of study design in which information on exposure and outcome are collection within the same time period[45]. In our study, we collected information on stress from baseline questionnaires and data on semen quality within a few weeks after baseline completion. No follow up is needed in cross-sectional studies. As data is collected at the same time, cross-sectional studies cannot evaluate causality, which requires a time separation between exposure and outcome. Thus, we cannot evaluate the causality between stress and semen quality. The study describes the prevalence of impaired semen quality among pregnancy planners in SF and PRESTO. However, it does not describe the prevalence of impaired semen quality among all men of reproductive age. It requires another sample, which is representative for men of reproductive age.

Missing data

In our study population, the proportion of missing data on perceived stress ranged from 1.9-5.8% in SF, while each PSS item in PRESTO had 0.36% missing (Supplementary table 4). Missing data is ubiquitous in clinical research and are often categorized into missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR). Data, which is MCAR, is independent of observed and unobserved data, whereas data MAR depends on the observed data[46]. For example, in our study data is MAR if well-educated men are more likely to report on the PSS, whereas data is MNAR if men with a higher stress level are less likely to report on the PSS. The proportion of men (24.4%) with missing data on sperm motility is MAR, as they pilot tested their semen quality before Trak® was able to measure sperm motility.

Multiple imputation is a statistical method to deal with missing data. In multiple imputation methods, the missing values are imputed based on the distribution of other variables in the datasets. It is often used under the assumption of data being MAR. The statistical method generates several complete datasets with plausible estimates of the missing values[34,46]. We generated five imputed datasets, analyzed each dataset separately and subsequently combined the results across the imputed datasets to

evaluate the association between perceived stress and semen quality. Missing motility data is not yet imputed.

Random and systematic error

Like any other epidemiologic study, we must consider if our findings between perceived stress and in-home assessed semen quality can be explained be systematic or random error.

Errors, which appear after removing systematic errors, are described as random error. Random error is the variability in dataset, which can be reduced by increasing the size of the study population[45]. We used 95% CI to describe the extent of random error or statistical variation related to the estimates. Findings from the linear regression provided estimates with narrow 95% CI, which indicate high precision. Estimates based on the logistic regression had wide 95% CI indicating low and inadequate precision. We treated the exposure and outcome as categorical variables in our logistic regression model. We dichotomized PSS (<20 vs. ≥20) and used WHO cut-off values for semen quality[42]. Due to a small study population it is difficult to ensure a sufficient number of subjects in all categories, which may explain why we obtain wide CI.

Errors which remain despite an increase in the study population are called systematic error. They are also referred to as bias[45]. They cause the estimates to systematically differ from the true value. Systematic errors result from selection bias, information bias and confounding.

Selection bias

Selection bias occurs when the association between exposure and outcome differs between those who participate in the study and those who do not participate. It can be caused by the way participants are selected or factors that influence study participation[45]. Our study is based on data from two prospective cohort studies of couples trying to conceive. Both cohorts enroll couples with fertility ranging from high to infertile. Couples with unintended pregnancies are not included in the cohort, as they tend to occur among the most fertile couples. This may cause both cohorts to overrepresent sub fertile couples. Men in our study may have a slightly lower semen quality compared to men in the general population. However, it is often the healthiest men and women who enroll in studies, which may cause the cohorts to represent a higher fertility than the general population. Thus, SF and PRESTO may be representative for the spectrum of fertility in the general population.

Internet access may influence enrollment, although it seems unlikely that internet access should differ according to participants and non-participants.

Information bias

Erroneous collection of information on exposure, covariates and outcome leads to information bias, which results in misclassification. Misclassification of study participants can be divided into differential

and non-differential misclassification. Differential misclassification appears when the misclassification differs according to other variables e.g. the outcome of interest. If the misclassification of a variable is unrelated to other variables, it is referred to as non-differential[45]. Differential misclassification may over- or underestimate the association. Non-differential misclassification of a non-dichotomous exposure may bias the estimate either towards or away from the null value, whereas non-differential misclassification of a dichotomous exposure may affect the estimate towards the null value.

We collected information on perceived stress prior to inviting men for in-home semen testing, thus potential misclassification should be non-differential. It could have been differential, if they were aware of their status of semen quality when reporting levels of perceived stress at baseline.

Recall bias is another type of information bias. It is present if the outcome of interest occurs prior to collecting data on the exposure. It leads to differential misclassification. Recall bias is irrelevant in our study as all data is collected within a short time period.

Confounding

Confounding can be defined as a confusion of effects[45]. It appears when the effect of the exposure on the outcome is mixed with the effect of another variable on the outcome. A variable is a confounder if it is associated with *both* the exposure and the outcome and it is imbalanced across exposure groups. The confounding variable should not be part of the causal pathway between the exposure and the outcome. Some studies control for confounding by methods within study design (randomization, matching and restriction), but it can also be done in the statistical analysis (stratification, adjustments or standardization). It reduces the amount of confounding but does not completely remove it. The remaining confounding is referred to as residual confounding.

We identified potential confounders based on our directed acyclic graph (DAG) (Figure 2) and existing literature. According to the DAG, cohort, age, BMI, physical activity, alcohol consumption, smoking and education are potential confounders. We adjusted for all variables except physical activity, which is a possible source of unmeasured confounding. It was excluded from the analysis due to technical reasons. Based on the DAG, sleep duration, diagnosis of anxiety or depression and abstinence time are intermediates. An intermediate is part of the causal pathway. Adjusting for an intermediate removes part of the association between the exposure and the outcome. To evaluate the effect from sleep duration, diagnosis of depression and anxiety, we adjusted for these variables in a sub-analysis (Table 2, model 2). Abstinence time is not considered a confounder, but we adjusted for the variable as it is strongly associated with semen quality. Some variables may not have been taken into account because they were unknown (e.g. fever episodes at sample collection). It may have led to unmeasured confounding.

References

1 Thoma, M. E. *et al.* Prevalence of infertility in the United States as estimated by the current duration approach and a traditional constructed approach. *Fertility and sterility* **99**, 1324-1331.e1321, doi:10.1016/j.fertnstert.2012.11.037 (2013).

2 Irvine, D. S. Epidemiology and aetiology of male infertility. *Human reproduction (Oxford, England)* **13 Suppl 1**, 33-44 (1998).

3 Levine, H. *et al.* Temporal trends in sperm count: a systematic review and meta-regression analysis. *Human reproduction update* **23**, 646-659, doi:10.1093/humupd/dmx022 (2017).

4 Priskorn, L. *et al.* Average sperm count remains unchanged despite reduction in maternal smoking: results from a large cross-sectional study with annual investigations over 21 years. *Human reproduction (Oxford, England)* **33**, 998-1008, doi:10.1093/humrep/dey090 (2018).

5 Ramlau-Hansen, C. H. *et al.* Is smoking a risk factor for decreased semen quality? A cross-sectional analysis. *Human reproduction (Oxford, England)* **22**, 188-196, doi:10.1093/humrep/del364 (2007).

6 Andersen, J. M. *et al.* Body Mass Index Is Associated with Impaired Semen Characteristics and Reduced Levels of Anti-Mullerian Hormone across a Wide Weight Range. *PloS one* **10**, e0130210, doi:10.1371/journal.pone.0130210 (2015).

7 Jensen, T. K. *et al.* Habitual alcohol consumption associated with reduced semen quality and changes in reproductive hormones; a cross-sectional study among 1221 young Danish men. *BMJ open* **4**, e005462, doi:10.1136/bmjopen-2014-005462 (2014).

8 Wise, L. A. *et al.* Male sleep duration and fecundability in a North American preconception cohort study. *Fertility and sterility* **109**, 453-459, doi:10.1016/j.fertnstert.2017.11.037 (2018).

9 Sundhedsstyrelsen. *Danskernes Sundhed - Den Nationale Sundhedsprofil 2017*, <https://www.sst.dk/-/media/Udgivelser/2018/Den-Nationale-Sundhedsprofil-2017.ashx?la=da&hash=421C482AEDC718D3B4846FC5E2B0EED2725AF517> (2018).

10 Witt, W. P. *et al.* Maternal stressful life events prior to conception and the impact on infant birth weight in the United States. *Am J Public Health* **104 Suppl 1**, S81-89, doi:10.2105/ajph.2013.301544 (2014).

11 Canada, S. Stress and well-being *Health Reports* **12**, 21-32 (2001).

12 Rosengren, A. *et al.* Association of psychosocial risk factors with risk of acute myocardial infarction in 11119 cases and 13648 controls from 52 countries (the INTERHEART study): case-control study. *Lancet* **364**, 953-962, doi:10.1016/s0140-6736(04)17019-0 (2004).

13 Liu, M. Y., Li, N., Li, W. A. & Khan, H. Association between psychosocial stress and hypertension: a systematic review and meta-analysis. *Neurol Res* **39**, 573-580, doi:10.1080/01616412.2017.1317904 (2017).

14 Rosmond, R., Dallman, M. F. & Bjorntorp, P. Stress-related cortisol secretion in men: relationships with abdominal obesity and endocrine, metabolic and hemodynamic abnormalities. *J Clin Endocrinol Metab* **83**, 1853-1859, doi:10.1210/jcem.83.6.4843 (1998).

15 Cohen, S., Janicki-Deverts, D. & Miller, G. E. Psychological Stress and Disease. *JAMA* **298**, 1685-1687, doi:10.1001/jama.298.14.1685 (2007).

16 Wesselink, A. K. *et al.* Perceived Stress and Fecundability: A Preconception Cohort Study of North American Couples. *American journal of epidemiology* **187**, 2662-2671, doi:10.1093/aje/kwy186 (2018).

17 Nordkap, L. *et al.* Psychological stress and testicular function: a cross-sectional study of 1,215 Danish men. *Fertility and sterility* **105**, 174-187.e171-172, doi:10.1016/j.fertnstert.2015.09.016 (2016).

18 Hjollund, N. H., Bonde, J. P., Henriksen, T. B., Giwercman, A. & Olsen, J. Reproductive effects of male psychologic stress. *Epidemiology (Cambridge, Mass.)* **15**, 21-27, doi:10.1097/01.ede.0000100289.82156.8b (2004).

19 Janevic, T. *et al.* Effects of work and life stress on semen quality. *Fertility and sterility* **102**, 530-538, doi:10.1016/j.fertnstert.2014.04.021 (2014).

20 Nargund, V. H. Effects of psychological stress on male fertility. *Nat Rev Urol* **12**, 373-382, doi:10.1038/nrurol.2015.112 (2015).
21 Byun, J. S. *et al.* Sexual dysfunctions induced by stress of timed intercourse and medical treatment. *BJU Int* **111**, E227-234, doi:10.1111/j.1464-410X.2012.11577.x (2013).
22 Eskiocak, S. *et al.* Effect of psychological stress on the L-arginine-nitric oxide pathway and semen quality. *Brazilian Journal of Medical and Biological Research* **39**, 581-588, doi:10.1590/S0100-879X2006000500003 (2006).
23 Schaff, U. Y. *et al.* Novel centrifugal technology for measuring sperm concentration in the home. *Fertility and sterility* **107**, 358-364.e354, doi:10.1016/j.fertnstert.2016.10.025 (2017).
24 Mikkelsen, E. M. *et al.* Cohort profile: the Danish Web-based Pregnancy Planning Study--'Snart-Gravid'. *International journal of epidemiology* **38**, 938-943, doi:10.1093/ije/dyn191 (2009).
25 Huybrechts, K. F. *et al.* A successful implementation of e-epidemiology: the Danish pregnancy planning study 'Snart-Gravid'. *European journal of epidemiology* **25**, 297-304, doi:10.1007/s10654-010-9431-y (2010).
26 Christensen, T. *et al.* Costs and Efficiency of Online and Offline Recruitment Methods: A Web-Based Cohort Study. *Journal of medical Internet research* **19**, e58, doi:10.2196/jmir.6716 (2017).
27 Hatch, E. E. *et al.* Evaluation of Selection Bias in an Internet-based Study of Pregnancy Planners. *Epidemiology (Cambridge, Mass.)* **27**, 98-104, doi:10.1097/ede.0000000000000400 (2016).
28 Wise, L. A. *et al.* Design and Conduct of an Internet-Based Preconception Cohort Study in North America: Pregnancy Study Online. *Paediatric and perinatal epidemiology* **29**, 360-371, doi:10.1111/ppe.12201 (2015).
29 Cohen, S., Kamarck, T. & Mermelstein, R. A global measure of perceived stress. *J Health Soc Behav* **24**, 385-396 (1983).
30 Lee, E. H. Review of the psychometric evidence of the perceived stress scale. *Asian Nurs Res (Korean Soc Nurs Sci)* **6**, 121-127, doi:10.1016/j.anr.2012.08.004 (2012).
31 Cohen, S. & M., W. G. in *The social psychology of health* (eds Stuart Oskamp & Shirlynn Spacapan) Ch. 3, 31-67 (Sage, 1988).
32 Sommer, G. J. *et al.* At-home sperm testing for epidemiologic studies: Evaluation of the Trak male fertility testing system in an internet-based preconception cohort. *Paediatric and perinatal epidemiology*, doi:10.1111/ppe.12612 (2019).
33 Kreiberg, M. *et al.* Psychological stress in long-term testicular cancer survivors: a Danish nationwide cohort study. *J Cancer Surviv*, doi:10.1007/s11764-019-00835-0 (2019).
34 Pedersen, A. B. *et al.* Missing data and multiple imputation in clinical epidemiological research. *Clin Epidemiol* **9**, 157-166, doi:10.2147/clep.S129785 (2017).
35 Organization, W. H. *WHO laboratory manual for the Examination and processing of human semen* <https://apps.who.int/iris/bitstream/handle/10665/44261/9789241547789_eng.pdf;jsessionid=8BD2C3D7BEE0E36F491A0B21AC48CA4F?sequence=1> (2010).
36 Ilacqua, A., Izzo, G., Emerenziani, G. P., Baldari, C. & Aversa, A. Lifestyle and fertility: the influence of stress and quality of life on male fertility. *Reproductive biology and endocrinology : RB&E* **16**, 115, doi:10.1186/s12958-018-0436-9 (2018).
37 Komité, N. V. *Om komitésystemet*, <http://www.nvk.dk/om-nvk/om-komitesystemet> (
38 Komité, N. V. *"Hvad skal jeg anmelde?"*, <http://www.nvk.dk/forsker/naar-du-anmelder/hvilke-projekter-skal-jeg-anmelde> (
39 Komité, N. V. *Før du beslutter dig*

Om at være forsøgsperson i sundhedsvidenskabelige forsøg, <http://www.nvk.dk/~/media/NVK/Dokumenter/Fr_du_beslutter_digpdf.pdf?la=da> (Revised March 11th 2019).
40 Wilcox, A. J. Ch. 9 Fertility and Fecundability, 123-141 (Oxford University Press, Incorporated, 2010).

41 Emile, L. B. a. (eds Emile L. Boulpaep & Walter F. Boron) Ch. 54, 1092-1107 (Elsevier, 2017).
42 Cooper, T. G. *et al.* World Health Organization reference values for human semen characteristics. *Human reproduction update* **16**, 231-245, doi:10.1093/humupd/dmp048 (2010).
43 Sundhedsstyrelsen. *Stress i Danmark - hvad ved vi?*, <https://www.sst.dk/da/udgivelser/2007//-/media/Udgivelser/2007/Publ2007/CFF/Stress/Stress_i_DK_11dec07,-d-,pdf.ashx> (December 2007).
44 Robert H. Fletcher, S. W. F., Grant S. Fletcher. (ed Fifth Edition) Ch. 5, 61-79 (Wolters Kluwer. Lippincott Williams & Wilkins, 2014).
45 Rothman, K. J. *Epidemiology - an introduction*. (Oxford University Press 2012).
46 White, I. R., Royston, P. & Wood, A. M. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* **30**, 377-399, doi:10.1002/sim.4067 (2011).

Supplementary tables

Supplementary table 1. Descriptive statistics of PSS score and semen quality				
	N	**Median (IQR)**	**10% percentile**	**90% percentile**
PSS score	328	14.0 (10.0-18.0)	7.0	23.0
Semen volume (ml)	328	3.8 (3.0-4.8)	2.1	6.0
Sperm concentration (million/ml)	328	49.0 (27.9-81.8)	14.5	125.0
Total sperm count (million)	328	191.1 (99.2-300.8)	42.8	467.1
Sperm motility (%)	248*	50 (40-70)	0.2	0.8

*80 men in PRESTO have missing data on sperm motility

Supplementary table 2. No. of men below WHO's lower reference limits for impaired semen quality			
	PRESTO	**SF**	**Total**
	N (%)	**N (%)**	**N (%)**
Semen volume <1.5 ml	5 (1.8%)	0 (0%)	5 (1.5%)
Sperm concentration <15 million	25 (9.1%)	12 (23.1%)	37 (11.3%)
Total sperm count <39 million	20 (7.2%)	7 (13.5%)	27 (8.2%)
Sperm motility <40%	62 (22.5%)*	17 (32.7%)	81 (32.66%)*

*80 men in PRESTO have missing data on sperm motility

Supplementary table 3. Participation rates in SF/Saedkvalitet					
Invited to participate	**Interested in participating**	**Completed phone call**	**Consented to participate**	**Successfully uploaded test 1 results and completed survey**	**Successfully uploaded test 2 results**
136	62 (45.6%)	58 (42.6%)	55 (40.4%)	52 (94.5%)	49 (89.1%)

Supplementary table 4. Missing values and covariates according to study					
PRESTO			**SF**		
Covariate	**N**	**%**	**Covariate**	**N**	**%**
m_educ	5	1.8115942	VocationalTraining	1	1.9230769
m_walkexer	1	0.3623188	JobHoursPerWeek	15	28.846154
m_cignum	265	96.014493	HourSleep	1	1.9230769
m_jogswimraq	1	0.3623188	Cystitis	2	3.8461538
m_weightnow	1	0.3623188	Chlamydia	2	3.8461538
m_snusfreq	267	96.739130	GenitalHerpes	2	3.8461538
m_nicotinefreq	270	97.826087	HPV	3	5.7692308
m_marijuanafreq	225	81.521739	Depression	2	3.8461538
m_hpvvacc	1	0.3623188	Anxiety	2	3.8461538
m_cellbackpocket	2	0.7246377	Diabetes	2	3.8461538
m_cellshirtpocket	4	1.4492754	InfectionMaleReproductiveOrgans	1	1.9230769
m_cellsidepocket	4	1.4492754	Height	1	1.9230769
m_cellbeltcarrier	3	1.0869565	Waist	23	44.230769
m_unemp	4	1.4492754	bmi	1	1.9230769
m_student	4	1.4492754	RedWineNumber	1	1.9230769
m_unempnw	4	1.4492754	WhiteWineNumber	2	3.8461538
m_home	4	1.4492754	DessertWineNumber	4	7.6923077
m_jobhrsperweek	15	5.4347826	LiquourNumber	5	9.6153846
m_pastsmoke	11	3.9855072	CoffeeNumber	1	1.9230769
m_weight17	22	7.9710145	DecafCoffeeNumber	1	1.9230769
m_preterm	2	0.7246377	BlackTeaNumber	1	1.9230769

m_goyourway	1	0.3623188	GreenTeaNumber	1	1.9230769
m_upset	1	0.3623188	WhiteTeaNumber	2	3.8461538
m_unablecontrol	1	0.3623188	HerbalTeaNumber	1	1.9230769
m_stressed	1	0.3623188	CokeNumber	1	1.9230769
m_handleproblems	1	0.3623188	CokeLightNumber	1	1.9230769
m_cope	1	0.3623188	Packyears	2	3.8461538
m_controlirritations	1	0.3623188	FeelingNervousStressed	2	3.8461538
m_topopthings	1	0.3623188	FeelingGoingYourWay	2	3.8461538
m_angrycontrol	1	0.3623188	FeelingNotCoping	2	3.8461538
m_difficultiespiling	1	0.3623188	FeelingControlIrritations	1	1.9230769
m_rappetite	1	0.3623188	FeelingOnTop	1	1.9230769
m_iappetite	1	0.3623188	FeelingAnger	1	1.9230769
m_troublesleeping	1	0.3623188	FeelingDifficulties	1	1.9230769
m_feltsad	1	0.3623188	FeelingSad	1	1.9230769
m_lostinterest	1	0.3623188	FeelingLostInterest	3	5.7692308
m_lackenergy	1	0.3623188	FeelingLackEnergy	2	3.8461538
m_lessconfident	1	0.3623188	FeelingLessSelfConfident	1	1.9230769
m_guilt	1	0.3623188	FeelingBadConscience	2	3.8461538
m_diffconcentrating	1	0.3623188	FeelingNotWorthLiving	2	3.8461538
m_restless	1	0.3623188	FeelingDifficultyConcentrating	1	1.9230769
m_wasntworthliving	1	0.3623188	FeelingRestless	1	1.9230769
m_subdued	1	0.3623188	FeelingSubdued	1	1.9230769
m_yoga	3	1.0869565	FeelingTroubleSleeping	1	1.9230769
m_esmk1120home	2	0.7246377	FeelingReducedAppetite	1	1.9230769
m_esmk2130home	2	0.7246377	FeelingIncreasedAppetite	1	1.9230769
m_waist	73	26.449275			
m_esmk010home	2	0.7246377			
m_esmk2130work	2	0.7246377			
m_esmkcurrhome	2	0.7246377			
m_esmkcurrwork	2	0.7246377			
m_numsexpart	45	16.304348			
m_laptophrsday	111	40.217391			
m_performenhance	240	86.956522			
m_esmk3140home	2	0.7246377			
m_esmk3140work	2	0.7246377			
m_ecigever	38	13.768116			
m_bikeshorts	191	69.202899			
m_bikeseat	192	69.565217			
m_seatheatershrswk	202	73.188406			
m_agestopsmoke	233	84.420290			
m_partnereduc	123	44.565217			
m_depevermed	238	86.231884			
m_ecigmgnicotine	261	94.565217			
m_ecigagefirst	237	85.869565			
m_deprx	246	89.130434			
m_anxevermed	245	88.768116			
m_ecigml	264	95.652174			
m_anxrx	252	91.304348			
m_hpvvaccage	262	94.927536			

Supplementary figures

Supplementary figure 1. Photos of Trak® Male Fertility System

Supplementary figure 2. Enrolment in SnartForaeldre.dk/Saedkvalitet