



## University of Groningen

### Genotype scores predict drug efficacy in subtypes of female sexual interest/arousal disorder

Tuiten, Adriaan; Michiels, Frits; Böcker, Koen Be; Höhle, Daniël; van Honk, Jack; de Lange, Robert Pj; van Rooij, Kim; Kessels, Rob; Bloemers, Jos; Gerritsen, Jeroen

*Published in:*

Women's health (London, England)

*DOI:*

[10.1177/1745506518788970](https://doi.org/10.1177/1745506518788970)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2018

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Tuiten, A., Michiels, F., Böcker, K. B., Höhle, D., van Honk, J., de Lange, R. P., ... Pfaus, J. G. (2018). Genotype scores predict drug efficacy in subtypes of female sexual interest/arousal disorder: A double-blind, randomized, placebo-controlled cross-over trial. *Women's health (London, England)*, 14. <https://doi.org/10.1177/1745506518788970>

#### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

#### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Genotype scores predict drug efficacy in subtypes of female sexual interest/arousal disorder: A double-blind, randomized, placebo-controlled cross-over trial

Women's Health  
Volume 14: 1–11  
© The Author(s) 2018  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/1745506518788970  
journals.sagepub.com/home/whe



**Adriaan Tuiten<sup>1</sup>, Frits Michiels<sup>2</sup>, Koen BE Böcker<sup>3</sup>, Daniël Höhle<sup>3</sup>, Jack van Honk<sup>4,5,6</sup>, Robert PJ de Lange<sup>3</sup>, Kim van Rooij<sup>1,7</sup>, Rob Kessels<sup>1</sup>, Jos Bloemers<sup>1,7</sup>, Jeroen Gerritsen<sup>1,7</sup>, Paddy Janssen<sup>8,9</sup>, Leo de Leede<sup>10</sup>, John-Jules Meyer<sup>3,11</sup>, Walter Everaerd<sup>12</sup>, Henderik W Frijlink<sup>13</sup>, Hans PF Koppeschaar<sup>1</sup>, Berend Olivier<sup>7,14,15</sup> and James G Pfaus<sup>16</sup>**

## Abstract

Attempts to develop a drug treatment for female sexual interest/arousal disorder have so far been guided by the principle of 'one size fits all', and have failed to acknowledge the complexity of female sexuality. Guided by personalized medicine, we designed two on-demand drugs targeting two distinct hypothesized causal mechanisms for this sexual disorder. The objective of this study was to design and test a novel procedure, based on genotyping, that predicts which of the two on-demand drugs will yield a positive treatment response. In a double-blind, randomized, placebo-controlled cross-over experiment, 139 women with female sexual interest/arousal disorder received three different on-demand drug-combination treatments during three 2-week periods: testosterone 0.5 mg + sildenafil 50 mg, testosterone 0.5 mg + buspirone 10 mg, and matching placebo. The primary endpoint was change in satisfactory sexual events. Subjects' genetic profile was assessed using a microarray chip that measures 300,000 single-nucleotide polymorphisms. A preselection of single-nucleotide polymorphisms associated with genes that are shown to be involved in sexual behaviour were combined into a Phenotype Prediction Score. The Phenotype Prediction Score demarcation formula was developed and subsequently validated on separate data sets. Prediction of drug-responders with the Phenotype Prediction Score demarcation formula gave large effect sizes ( $d=0.66$  through  $1.06$ ) in the true drug-responders, and medium effect sizes

<sup>1</sup>Emotional Brain BV, Almere, The Netherlands

<sup>2</sup>Chemistry and Life Sciences, V.O. Patents & Trademarks, Amsterdam, The Netherlands

<sup>3</sup>Alan Turing Institute Almere, Almere, The Netherlands

<sup>4</sup>Department of Experimental Psychology, Utrecht University, Utrecht, The Netherlands

<sup>5</sup>Institute of Infectious Disease and Molecular Medicine (IDM), University of Cape Town, Cape Town, South Africa

<sup>6</sup>Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa

<sup>7</sup>Utrecht Institute for Pharmaceutical Sciences and Rudolf Magnus Institute of Neuroscience, Utrecht University, Utrecht, The Netherlands

<sup>8</sup>Division of Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands

<sup>9</sup>Department of Central Hospital Pharmacy, Viecuri Hospital, Venlo, The Netherlands

<sup>10</sup>Exelion Bio-Pharmaceutical Consultancy B.V., Waddinxveen, The Netherlands

<sup>11</sup>Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands

<sup>12</sup>Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

<sup>13</sup>Research Group of Pharmaceutical Technology and Biopharmacy, University of Groningen, Groningen, The Netherlands

<sup>14</sup>Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA

<sup>15</sup>Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands

<sup>16</sup>Department of Psychology, Centre for Studies in Behavioral Neurobiology, Concordia University, Montreal, QC, Canada

## Corresponding author:

Adriaan Tuiten, Emotional Brain BV, Louis Armstrongweg 78, 1311 RL, Almere, The Netherlands.  
Email: a.tuiten@emotionalbrain.nl



( $d=0.51$  and  $d=0.47$ ) in all patients (including identified double, and non-responders). Accuracy, sensitivity, specificity, positive predictive value, and negative predictive value of the Phenotype Prediction Score demarcation formula were all between 0.78 and 0.79, and thus sufficient. The resulting Phenotype Prediction Score was validated and shown to effectively and reliably predict which women would benefit from which on-demand drug, and could therefore also be useful in clinical practice, as a companion diagnostic establishing the way to a true personalized medicine approach.

### Keywords

female sexual interest/arousal disorder, genotype scores, hypoactive sexual desire disorder, personalized medicine, phenotype prediction score, satisfactory sexual events, single-nucleotide polymorphisms, testosterone

Date received: 22 March 2018; revised: 17 May 2018; accepted: 25 June 2018

Low sexual desire and/or arousal are the most common sex-related complaints reported by women.<sup>1,2</sup> They often result in sexual dissatisfaction, which in turn impacts psychological well-being and can result in severe personal distress.<sup>3</sup> These complaints are classified in the Diagnostic and Statistical Manual for Mental Disorders, edition 5 (*DSM-5*)<sup>4</sup> as female sexual interest/arousal disorder (FSIAD). The disorder is likely caused by a complex interaction of psychological and neurobiological factors<sup>5-7</sup> and is prevalent among women of all ages<sup>8</sup> and ethnicities.<sup>9</sup> While effective pharmacological treatments for erectile dysfunction (ED) have been available for two decades now,<sup>10</sup> the development of effective and safe drug treatments for sexual dysfunction in women has met with only limited success. Flibanserin recently became available for treatment of premenopausal women with FSIAD<sup>11</sup> in the United States, and later in Canada, but it shows poor efficacy and produces a range of adverse effects. This drug requires the use of an extensive risk evaluation and mitigation strategy to prevent untoward adverse effects. Moreover, no drug treatment is yet available outside North America.

In seeking effective treatment for FSIAD, we have adopted a hypothesis-driven, personalized sexual medicine approach,<sup>12-15</sup> based on established neurobiological mechanisms known to either promote or inhibit sexual desire.<sup>5</sup> We have identified a subgroup of women with FSIAD in which the brain systems for sexual excitation are relatively insensitive, resulting in low sexual desire/arousal. In addition, we have identified another substantial subgroup of women in which exposure to sexual stimulation dysfunctionally increases the activity of inhibitory sexual mechanisms. This, too, results in low sexual desire/arousal. The neurobiological mechanism involved in this inhibitory effect likely involves a phasic increase in serotonergic activity in the left dorsolateral prefrontal cortex (DLPFC),<sup>5,12</sup> which is elicited by sexual stimulation. This subgroup division is based on the dual-control model of sexual response<sup>16</sup> and is substantiated by cognitive,<sup>15,17</sup> psychophysiological,<sup>13-15,17,18</sup> subjective,<sup>13-15,19</sup> neuroanatomical,<sup>12,20</sup> and pharmacological evidence.<sup>13-15,17,19</sup> Unfortunately, it is not readily apparent to which category a patient belongs because the symptoms,

low desire and/or low arousal, manifest in the same manner in both these subgroups. To summarize, it is believed that while low sensitivity to sexual cues is associated with low sexual desire and arousal, high sensitivity to sexual cues can also result in low desire/arousal if sexual stimulation elicits a dysfunctionally high sexual inhibition response.<sup>5,12</sup>

Accordingly, we have developed two different drug treatments targeted at these two distinct neurobiological mechanisms. The administration of a single dose of 0.5 mg sublingual testosterone (T) produces a peak in plasma level of T within 15 minutes, with a return to baseline within 2 hours. However, in a period ranging from 3 to 6 hours after peak plasma levels, sublingual T produces an increase in vaginal arousal and in subjective sexual responses in sexually functional women.<sup>21,22</sup> The developed treatments consist of novel (dual-route/dual-release, fixed-dose) combination tablets consisting of a T coating for sublingual administration, and an inner-core component containing the phosphodiesterase type 5 (PDE-5) inhibitor sildenafil (S) or the 5HT<sub>1A</sub> receptor agonist buspirone (B). The inner-core component is coated with a delayed-release matrix to ensure that the peak plasma concentration of S or B, respectively, coincides with the window of increased sexual motivation induced by the sublingual T. T + S has been developed for women with FSIAD and a relative insensitivity to sexual cues, and T + B for women with FSIAD and a dysfunctional activation of sexual inhibitory mechanisms. Earlier proof of concept experiments showed that both combinations (T + S and T + B) are promising potential on-demand treatments for two subtypes of women with FSIAD.<sup>13,14</sup> A description and explanation of the working mechanisms of the two drugs can be found in the Supplementary information.

To predict which of the two drug treatments will be effective for a woman diagnosed with FSIAD, we have searched for (combinations of) biological and psychological markers related to either one or both of the two distinct neurobiological mechanisms underlying FSIAD. An emotional Stroop task<sup>13,15,17</sup> and a combination of questionnaires and biological markers<sup>19</sup> were to a sufficient extent effective in separating T + S and T + B responders; however, this solution was deemed impractical for the clinical setting.

Therefore, we searched for combinations of genetic markers that reflect the aetiology of the phenotype and thus predict the responses to these drugs. We selected genetic markers from candidate genes associated with neurobiological mechanisms involved in the regulation of sexual behaviour. We combined the contribution of the selected genetic markers into a single score that predicts the phenotypes: the Phenotype Prediction Score (PPS).

There were four reasons for not applying an approach in which single-nucleotide polymorphisms (SNPs) for the PPS are selected based on a hypothesis-free genome wide association study (GWAS)<sup>23–25</sup> First, it would discard the current neurobiological knowledge of female sexual behaviour from our hypothesis-driven approach. Second, it would require thousands of women diagnosed with FSIAD, treated with both drugs, to obtain sufficient statistical power. Third, in recent (GWAS) studies of low sexual desire/arousal in women, only a few weak associations (oestrogen 2 receptor SNP rs4986938 was nominally significantly associated with both desire and lubrication before but not after correction for multiple testing)<sup>26</sup> or no significant associations<sup>27</sup> were found. Fourth, the present research is focussed on only one symptom, low sexual desire/arousal, for which we assume that different causal neurobiological mechanisms are responsible. From a logical point of view, it would be impossible to find a risk SNP for one symptom with different (opposing) causal mechanisms among individuals.

The low penetrance of individual genetic loci often means that no single genotype could determine or predict a phenotype.<sup>28,29</sup> Accordingly, we adopted an alternative approach that was more suited to FSIAD's complex nature and its phenotypes. Fisher<sup>30–32</sup> noted that traits influenced by multiple genes will have normal distributions, as quantitative traits. Thus, complex disorders affected by multiple genes, such as FSIAD, can be studied as quantitative traits.<sup>33</sup> Some individuals carry few genetic risk factors while others have many. The majority of the population will be clustered around the mean.

Our novel approach of the genetic subtyping of FSIAD patients, resulting in a PPS, was based on: (1) a selection of genes that affect neurochemical systems known to be involved in the regulation of female sexual behaviour<sup>5,26</sup> (see Supplementary information); (2) selection of SNPs associated with these genes (see for comparable hypothesis and genetic-driven selections of SNPs<sup>34–36</sup>); and (3) calculation of cumulative effect of multiple selected SNPs that emerged, resulting in the final PPS.

## Method

To develop an SNP-based subtyping system, we carried out a double-blind, randomized, placebo-controlled, crossover experiment in 139 women with FSIAD. After a 2-week baseline measurement, in which they received

placebo (single-blind placebo run-in period or PRI), the women were randomly assigned to different 2-week, double-blind, on-demand drug treatment periods. The primary endpoint in this experiment was an increase in the number of satisfactory sexual events (SSEs) between PRI and the active treatments (T + S; T + B). The experimental design included a within-comparison; therefore, the results for each participant consisted of two drug effects, one for T + S and one for T + B. For the purpose of our novel PPS procedure, we calculated a single-outcome measure, which incorporated both drug effects.

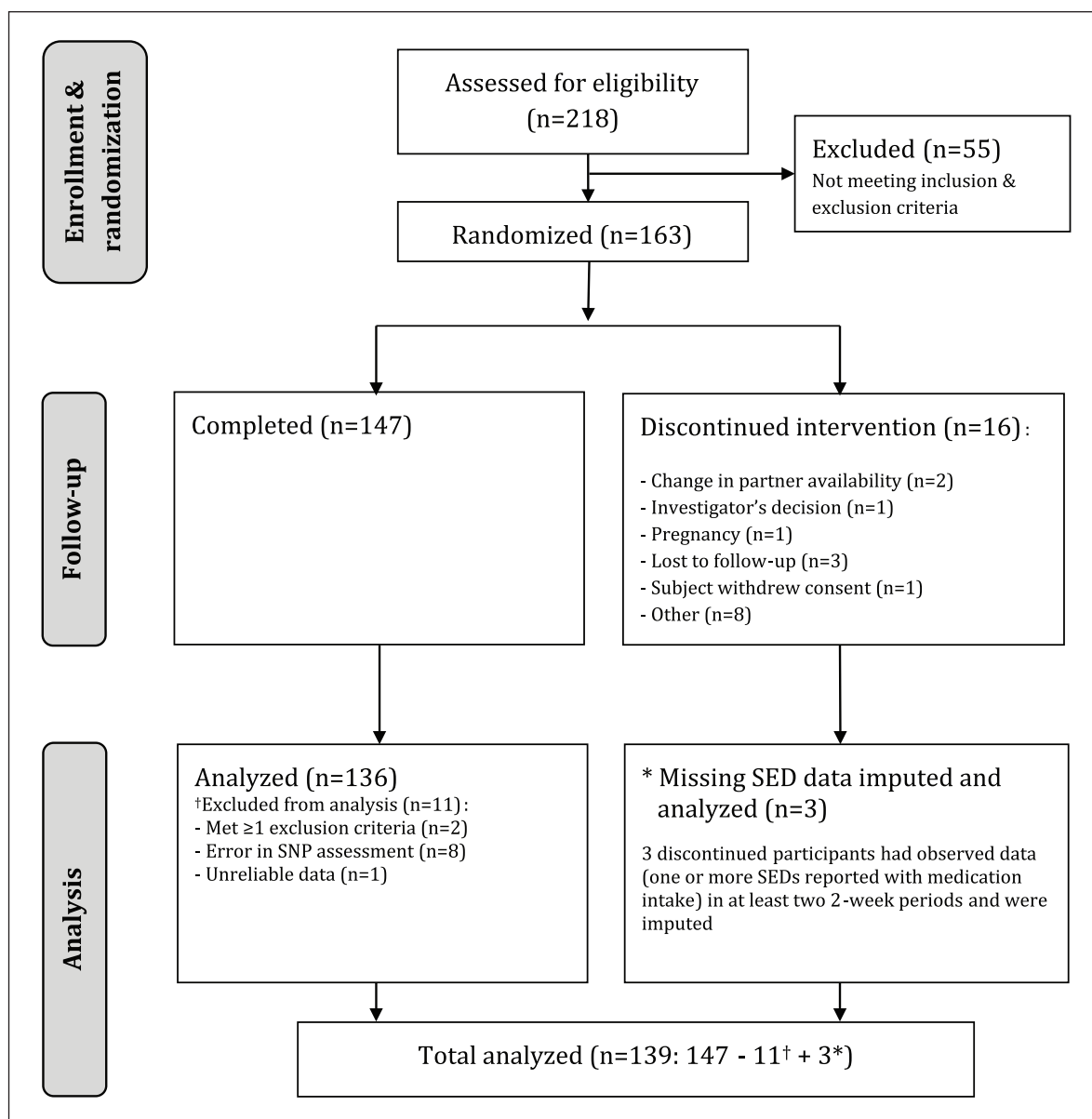
In developing the SNP-based subtyping system, we started by selecting a relatively small number of SNPs, based on a hypothesis-driven selection of genes involved in regulating sexual behaviour (see Supplementary information). The SNPs considered were those located inside these candidate genes or their promotor regions, or those located between a candidate gene and a neighbouring gene. From this selection, the SNPs with potential predictive value in terms of the response to each of both FSIAD treatments were distilled. These SNPs were combined into the PPS, creating a classification system that differentiates T + S and T + B responders based on a women's genetic profile. The PPS demarcation formula was developed in a derivation sample and then validated using an independent sample.

## Study aim

To develop and validate an SNP-based demarcation formula intended to predict the efficacy of two different drugs in two distinct subgroups of women with FSIAD. For this purpose, we used a clinical experiment to investigate the difference in efficacy of T + S (0.5 mg sublingual testosterone combined with 50 mg sildenafil) and T + B (0.5 mg sublingual testosterone combined with 10 mg buspirone) compared to placebo, in women with FSIAD.

## Study participants

Women between the ages of 18 and 70 were recruited via advertisements and from a database of volunteers (maintained by Emotional Brain BV, Almere, The Netherlands). Determination of the required sample size was done by a receiver operating characteristic (ROC) curve power analysis. Assuming a balanced design of 75 cases in each predicted group, 90% statistical power was required to detect a minimum area under the curve (AUC) of 0.65. After providing written informed consent, 218 women were screened for eligibility. Diagnoses of FSIAD were based on a semi-structured interview (see Supplementary information) that was conducted by two trained psychologists and were based on the *DSM-5*.<sup>4</sup> Individuals exhibiting comorbidity with female orgasmic disorder (as a secondary diagnosis) were not excluded. The subjects' medical



**Figure 1.** Consort diagram.

history was recorded, after which they were given a physical examination, including a 12-lead electrocardiogram and a urine pregnancy test. The standard biochemical, serological, and haematological laboratory parameters were assessed. Baseline levels of total testosterone (to rule out hyperandrogenism), sex hormone-binding globulin, albumin, thyroid stimulating hormone, follicle-stimulating hormone, luteinizing hormone, prolactin, and oestrogen, were also assessed at screening. The participants included in the final analyses were involved in a stable, communicative, monogamous relationship and had a sexually functional partner, had a normal medical history, and were otherwise healthy as determined by the physical examination, laboratory values, and vital signs

(see CONSORT diagram, Figure 1). Demographics of all randomized participants (N = 163) are presented in Table 1.

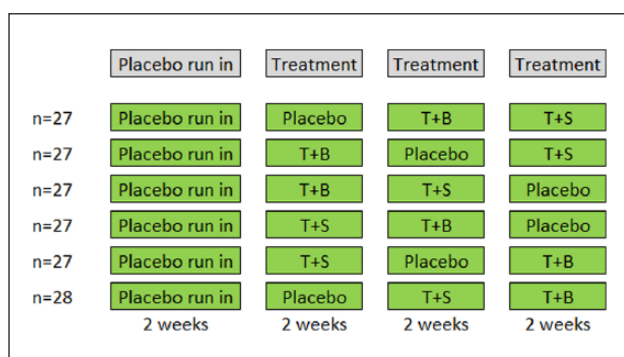
### Study performance and approval

Between 7 February 2014 (first screening visit) and 12 August 2014 (last follow-up visit, as per protocol), the study was conducted by trained research staff at two clinical research units of Emotional Brain BV (Almere and Utrecht) and monitored by PSR Group (Hoofddorp, The Netherlands). It was performed in compliance with the Declaration of Helsinki (October 2008) and with the International Conference on Harmonization – Good Clinical Practice guidelines for clinical research. The study

**Table 1.** Demographics all randomized participants.

Parameter	Number of participants (%) Total (N = 163)
Age (category)	
<40	105 (64.4)
40–60	50 (30.7)
≥60	8 (4.9)
Age (Years)	
Mean	34.7
Minimum	18.0
Maximum	67.0
Body mass index	
<35	160 (98.2)
≥35	3 (1.8)
Menopausal status	
Post-menopausal	28 (17.2)
Pre-menopausal	135 (82.8)

Denominator for the calculation of percentages: total number of participants randomized.

**Figure 2.** Study and randomization design.

was approved by a Medical Ethics Committee ‘Stichting BEBO’ (Assen, The Netherlands) and by the Dutch Competent Authority (Centrale Commissie Mensgebonden Onderzoek), under authorization number NL44803.056.13. It was registered in the European Clinical Trials Database, as EudraCT number 2011-000457-23. The trial was registered under Primary Registry trial number NTR4426 (Netherlands Trial Register).

### Study design

The study was a randomized, double-blind, placebo-controlled, cross-over study. It included a 2-week single-blind placebo run-in period (PRI), resembling baseline measurements during the phase 2 and planned phase 3 experiments, and three 2-week double-blind treatment periods (placebo, T + S, and T + B), with a 1-week follow-up period. Each regimen was separated by at least one 2-day wash-out period. (note that both drugs are intended for on-demand use and that systemic clearance takes

approximately 24 hours.) Each participant completed the three 2-week treatment periods in a randomized order, in accordance with a 6-sequence Williams design (see Figure 2). Treatment sequences were listed in a randomization list generated by independent statisticians at the Pharma Consulting Group (Uppsala, Sweden). Those who were eligible to participate in the study were randomized and allocated a treatment sequence by the principle investigator, using an interactive web-response system that was an integral part of the electronic case report form (Viedoc™, version 3.22; Pharma Consulting Group, Uppsala, Sweden). The unique 3-digit code allocated to each participant coincided with a unique 3-digit treatment code on each medication container. Prior to database lock, none of the study participants, none of the research staff involved, nor anyone employed by the sponsor had access to the randomization list.

Each of the participants visited the study site on a total of seven occasions. These consisted of one screening visit, one start-up visit, four study regimen follow-up visits, and one final follow-up visit. During the start-up visit and study regimen follow-up visits, an evaluation was made of each individual’s sexual functioning, their health was monitored, and study medication was dispensed. At the start-up visit, blood was drawn for the assessment of the SNPs.

The primary endpoint was the change in the total number of SSEs after medication intake between the single-blind placebo run-in period (PRI) and the double-blind active treatment period (ATP), as measured by the sexual event diary (SED). This is a more adequate endpoint for testing the efficacy of on-demand drugs than questionnaires assessing sexual functioning over longer periods of time (e.g. the preceding 4 weeks) as the observation is more proximate to an on-demand drug’s efficacy.

### SED

The SED is an 11-item, standardized quantitative scale, for which validated US-English<sup>37</sup> and Dutch<sup>38</sup> versions are available. Eight of these items (sexual satisfaction, sexual desire, physical arousal, mental arousal, sexual pleasure, orgasm, distraction, and inhibition) assess the individual’s sexual functioning during a single sexual event. The remaining three items assess the type and time of the sexual event, and whether the on-demand medication was used (as instructed). This questionnaire is filled out within 24 hours of the event. Only those events during which medication was used were analysed.

### Medication, dosing, and instructions

**Sublingual T + S.** It is a dual-route/dual-release fixed-dose combination of T and S citrate.<sup>39</sup> The drug product is a 9-mm, round, biconvex, white, menthol-flavoured tablet for sublingual administration. The outer coating (a polymeric film) contains testosterone (0.5 mg) that is released

immediately on sublingual administration. The inner core of the tablet, which contains sildenafil (50 mg), has a polymeric coating designed to delay the release of that drug for approximately 2.5 hours. When that period elapses, the sildenafil is released immediately (i.e. there is no sustained release).

**Sublingual T + B.** It is a dual-route/dual-release fixed-dose combination of T and B hydrochloride.<sup>40</sup> The drug product is a 9-mm, round, biconvex, white, menthol-flavoured tablet for sublingual and oral administration. The appearance, method of administration and flavour of T + B is identical to T + S. The outer, polymeric-film coating contains testosterone (0.5 mg) that is released immediately on sublingual administration. The inner core of the tablet contains buspirone hydrochloride (10 mg). This inner core has a polymeric coating designed to delay the release of buspirone for approximately 2.5 hours. When that period elapses, the buspirone is released immediately (i.e. there is no sustained release).

**Placebo.** The placebo tablets were identical, in terms of appearance and flavour, to the fixed-dose combination T + S and T + B tablets containing the active pharmaceutical ingredients. All of the medication used was manufactured and packaged at Piramal Healthcare UK (Morpeth, UK).

**Dosing instructions.** The participants were instructed to keep the tablet under their tongue for 90 seconds and then to swallow it whole, without chewing it or otherwise disrupting the dosage form. They were permitted to take the tablet with a little water. The subjects were instructed to engage in sexual activity between 3 and 6 hours after ingestion. The dosing instructions were the same for all medications.

**Duration of treatment.** A total of eight doses per regime were provided. The participants were asked to endeavour to take a minimum of four doses during the 2-week treatment periods (two doses/week). They were informed that they could take the remaining four doses as required (i.e. 'on demand') throughout the 2-week treatment period. The minimum period between individual doses was two days (i.e. on alternate days).

### SNP analysis

DNA extraction was performed by Medigenomix (Ebersberg, Germany). SNP analysis was performed using a microarray chip (HumanCytoSNP-12 bead chip, Illumina, containing 297,622 SNPs). The hybridization and chip readout were performed by Eurofins (The Netherlands), AROS Applied Biotechnology (Aarhus, Denmark), and Medigenomix (Ebersberg), specialized laboratory service providers.

### Statistical methods

**Missing data and imputation.** If participants failed to fill out an SED within a 2-week period, this resulted in missing data. This meant that either the participant did not experience a sexual event or that they did, but then forgot to report it. The percentage of missing SEDs during PRI was 2.9%, while during the placebo regime it was 6.5%. The corresponding values for the T + S regime and the T + B regime were 5.0% and 3.6%, respectively. Those participants who only reported SED in a single regime, or who reported no SEDs at all, were eliminated from the analysis. Those who reported SEDs in at least two regimes were imputed. The missing SEDs were imputed by first imputing the number of events involved (by sampling participants with similar response patterns). Then, the corresponding Likert-type scale items were imputed by the mean, based on participants with the same number of events.

**Statistical analyses.** Group-level statistics and patient-level statistics were derived to assess the efficacy, usefulness, and validity of the formula. Group-level analyses were performed using paired sample t-tests to assess the change from PRI to T + S and to T + B. This was assessed for the derivation sample (N=50), validation sample (N=47), and total sample (N=97) of observed responders. T + S response is defined as having more SSEs in the T + S than T + B condition. T + B response is defined as having more SSEs in the T + B than T + S condition. Furthermore, utility was tested for all participants, including observed non-responders (N=139). A multiplicity correction was applied by controlling the false discovery rate (FDR), to control for inflated Type 1 error rates. In addition to p-values, effect sizes (Cohen's d) were calculated using the formula for paired sample t-tests, as defined in Dunlop et al.<sup>41</sup>

Utility at patient-level was tested by deriving an ROC curve for the total sample (N=97) of observed responders. The AUC of the ROC curve served as the test statistic for the relationship between demarcation formula outcome and response status based on the outcome measure. The hypothesis tested was that, when used to classify women as either T + S or T + B responders, the formula would perform above chance level. The AUC can be interpreted as the probability that, for a given random pair (T + S responder plus a T + B responder), the formula would produce a larger outcome for the former than for the latter. ROCs with an AUC in excess of 0.714 indicated a large effect size ( $d' > 0.8$ ), while ROCs with an AUC in excess of 0.76 indicated an effect size of at least 1.<sup>42</sup> Finally, ROCs with an AUC in excess of 0.80 were generally considered to be 'good', and suitable for clinical use. Besides the ROC curve with corresponding AUC, classification performances were also calculated for the total sample of observed responders. These involved accuracy, sensitivity, specificity, positive predictive value (PPV),

**Table 2.** PPS SNP composition.

SNP identifier	Gene	Abbreviation	Implicated in
rs963468	Dopamine D3 receptor	DRD3	Wanting*
rs2770296	5-HT2a receptor	HTR2A	Novelty seeking
rs11168048	5-HT4 receptor	HTR4	Mediator of the neurogenic and behavioural actions of antidepressants*
rs3740046	5-HT7 receptor	HTR7	
rs140701	5-HT Transporter	SLC6A4	Alcohol intake behaviour, schizophrenia, panic disorder
rs13278849	Adrenoreceptor alpha 1A	ADRA1A	Olfactory-driven behaviours*
rs1079078	Adrenoreceptor alpha 1A	ADRA1A	Olfactory-driven behaviours*
rs10515805	Adrenoreceptor alpha 1B	ADRA1B	Lordosis*
rs12653825	Adrenoreceptor alpha 1B	ADRA1B	Lordosis*
rs41154	NE transporter	SLC6A2	
rs6259	SHGB		
rs7761133	Oestrogen 1 receptor	ESR1	
rs1256114	Oestrogen 2 receptor	ESR2	
rs7734558	Prolactin		
rs816353	Nitric oxide synthase	NOS1	Vasodilation
rs48255	Nitric oxide synthase	NOS1	Vasodilation

SNP: single-nucleotide polymorphism; SHGB: sex hormone-binding globulin; NE: norepinephrine.

See supplementary information for a full overview of these SNPs and their implicated function.

\*Gene implicated, SNP not described in the literature.

and negative predictive value (NPV). These classification performance statistics are usually used to obtain information about the ability of companion diagnostic tools to identify women with and without a given diagnosis. In this application, the tool was designed to classify women as either T + S or T + B responders, and these classification performances were interpreted accordingly. In this context, accuracy was reflected by the proportion of women classified correctly, while sensitivity corresponded to the formula's ability to identify T + B responders. Specificity was defined as the ability to identify T + S responders, PPV as the proportion of correctly identified T + B responders, and NPV as the proportion of correctly identified T + S responders.

## Results

The development of the PPS resulted in the combination of 16 SNPs (see Table 2).

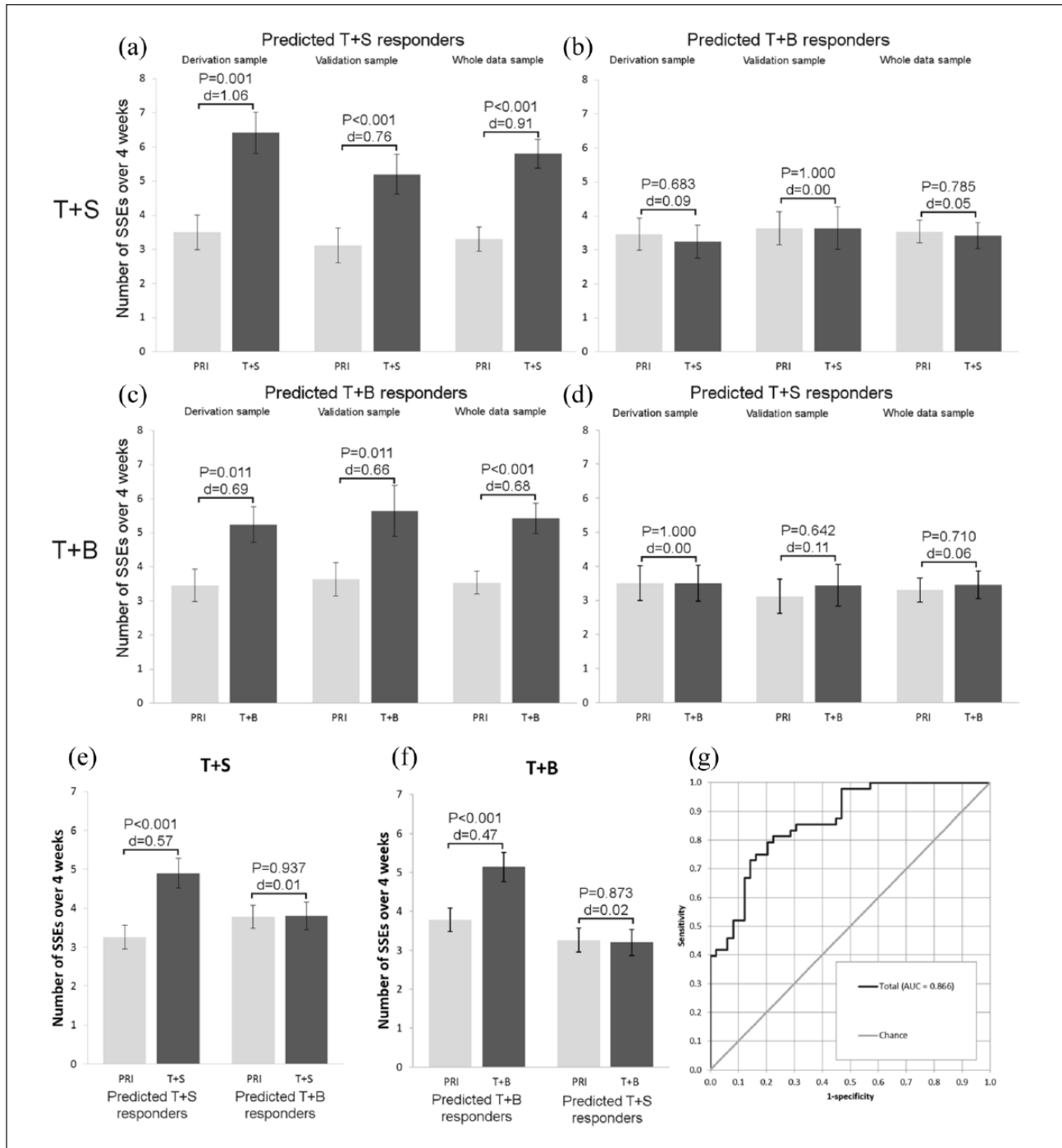
Figure 3 shows that the two FSIAD subtypes classified by the PPS demarcation formula (calculated as a dichotomy) exhibited clear improvements in symptoms in both the derivation and validation samples. This supports the validity of the PPS demarcation formula at the level of the responder group, which showed large effect sizes ( $d=0.66-1.06$ ). Moreover, the effects in all patients (including identified double, and non-responders) also showed the expected drug responses, and medium effect sizes ( $d=0.51$  and  $d=0.47$ ). Figure 3(g) shows the ROC curve of the PPS, which supports the utility of the PPS at the individual level as well.

T + S and T + B were well tolerated and no drug related serious adverse events were observed. None of the adverse events led to discontinuation of the study drug. Most adverse events were characterized as mild or moderate in severity and were consistent with the approved labelling for S, B and/or T. See Table 3 for the most common treatment-emergent adverse events per drug condition.

## Discussion

The idea behind the PPS method is that disorders like FSIAD result from an interplay between genetic background, past experiences, and present circumstances. Even though it is based solely of biological variables, it accounts for both biological and psychological mechanisms. For example, high sexual inhibition might result from the combination of genetic factors that increase the brain's sensitivity to sexual cues and adverse sexual experiences. A highly sensitive system to sexual stimuli combined with positive sexual experiences may lead to a pleasant and enjoyable sexual life. However, adverse sexual experiences may have a greater impact on those whose brains are more sensitive to sexual cues. This combination of a highly sensitive brain and adverse experiences may lead to a learned and autonomous sexual inhibitory response, in which sexual events and adverse associations are automatically linked. The concept of a dichotomous measure, which closely reflects a contrast in content-meaning, enabled us to identify risk genotypes that were linked to a series of SNPs and which may have opposing roles in terms of brain activity. This view reflects the possibility





**Figure 3.** (a, b, c, and d) The mean drug responses for responders in the derivation sample (N = 50), validation sample (N = 47), and total sample (N = 97) are shown for the primary endpoint of predicted T + S responders (N = 24, 25, and 49 for the derivation, validation, and total samples of responders, respectively; Panels a and d) and of predicted T + B responders (N = 26, 22, and 48 for the derivation, validation, and total samples, respectively; Panels b and c). Panels A and B are the results for the T + S responses ((a) T + S response for the predicted T + S responders, and (b) T + S response for the predicted T + B responders), while panels c and d are the results for the T + B responses ((c) T + B response for the predicted T + B responders and (d) T + B response for the predicted T + S responders). (e and f) The mean drug responses on the primary endpoint are shown for all participants (N = 139, including identified double-, and non-responders) in Figure 3, panels e and f. Panel e shows the response on the primary endpoint of T + S treatment for the predicted T + S responders (N = 67) and for the predicted T + B subgroup (N = 72). Panel f reveals the effect of treatment with T + B on the primary number of SSEs in the predicted T + S responders and the predicted T + B responders.

(Continued)

**Figure 3.** (Continued)

(a, b, c, d, e, and f) Each of the treatments was taken for 2 weeks, yet the data represent the mean number of SSEs over a 4-week period (to make it comparable with the results of similar experiments in this field). The reported p-values are two-sided. Error bars represent the standard error of the mean. To assess significance, p-values were tested against an alpha level of 0.05. A multiplicity correction was applied by controlling the false discovery rate (FDR) to control for inflated Type I error rates. As a result of controlling the FDR, all p-values below 0.025 were significant. Furthermore, the effect sizes were either medium ( $d \geq 0.50$ ) or large ( $d > 0.80$ ). Thus, the results with regard to the primary endpoint lead to the conclusion that the formula was adequately validated for all patients. Effect sizes were derived using the formula for calculating Cohen's d for paired sampled t-test.<sup>41</sup> (g). ROC curve for observed responders in the total sample (N=97). The p-value for the Area Under the Curve (AUC) was significant ( $p < 0.001$ ). Accuracy was 0.78. Sensitivity, defined as the ability to identify T + B responders, was 0.78. Specificity, defined as the ability to identify T + S responders, was 0.79. Positive predictive value, defined as the proportion of correctly classified T + B responders, was 0.79. Negative predictive value, defined as the proportion of correctly classified T + S responders, was 0.78. The combination of the nine inhibition SNPs and the seven B-coded SNPs resulted in a greater proportion of correctly classified patients. This combination was therefore superior to the nine inhibition SNPs alone. The ROC curve supports the usefulness of the formula at the individual level (Figure 3(g)). The AUC of the ROC curve served as the test statistic for the relationship between demarcation formula outcome and response status based on the outcome measure. The hypothesis tested was that, in classifying patients as either T + S or T + B responders, the formula would perform above chance level. The AUC was interpreted as the probability that, for a given random pair – T + S responder plus a T + B responder – the formula would produce a larger outcome for the former than for the latter. ROCs with an AUC in excess above 0.71 indicated a large effect size ( $d' > 0.8$ ), and ROCs with an AUC above 0.76 indicated an effect size of at least 1. Finally, ROCs with an AUC in excess of 0.80 were generally considered to be 'good' and suitable for clinical use.<sup>42</sup> The ROC curve showed an excellent AUC (Figure 3(g)).

SSE: satisfactory sexual event; d: effect size (Cohen's d); PRI: placebo run-in; T + S: testosterone 0.5 mg + sildenafil citrate 50 mg; T + B: testosterone 0.5 mg + buspirone hydrochloride 10 mg; ROC: receiver operator characteristic; AUC: area under the curve.

**Table 3.** Incidence of most common treatment-emergent adverse events per drug condition.

SOC	Term (MeDRA)	Study drug		
		Placebo	Lybrido	Lybridos
Gastrointestinal disorders	Nausea	2 (1.2)	8 (4.9)	13 (8)
Nervous system disorders	Dizziness	3 (1.8)	7 (4.3)	38 (23.3)
	Headache	10 (6.1)	28 (17.2)	10 (6.1)
Respiratory, thoracic and mediastinal disorders	Nasal congestion	1 (0.6)	10 (6.1)	1 (0.6)
Vascular disorders	Flushing	1 (0.6)	11 (6.7)	4 (2.5)
	Total	17	64	66

MeDRA: Medical Dictionary of Regulatory Activities; SOC: system organ class.

Adverse events that were summarized were reported after Visit 2 and before Visit 5 (or last intake investigational study drug).

Subjects with one or more adverse events within a level of the MeDRA term were counted only once for that level, except when study drug differed for the same event.

Most common is defined as  $\geq 5\%$  in any of the drug conditions.

Denominator for the calculation of percentages: total number of subjects randomized.

that opposing causal mechanisms could be responsible for low sexual desire/arousal.

This study has several limitations. The microarray chip used in the study contained 300,000 SNPs out of a possible 10 million, approximately. It is very well possible that another combination of SNPs exists that yield a better prediction. We have focussed on SNPs from a limited number of genes that have been associated with female sexual behaviour, in literature. It is, of course, conceivable that SNPs from other genes whose role in female sexual behaviour is still unclear, may improve the response prediction. Also, the study had a relatively low number of subjects. This increases the likelihood that the results are not fully generalizable to the full FSIAD population. The setup of the analyses however, using separate derivation and

validation sets, and the medium to large effect sizes, do increase the likelihood of this generalizability. Future research will need to establish if other SNPs/genes are even better predictors and if the results are generalizable. Moreover, it would be interesting to see whether this conditional genotyping method is also able to predict treatment outcome in other psychological/psychiatric disorders.

In conclusion, we have demonstrated that our conditional method that can be used to describe traits (i.e. differentiated treatment effects) partly in terms of unique sets of mutually exclusive risk genotypes. Moreover, combining the identified SNPs resulted in a PPS that could be used to predict correct drug responses in identified subtypes where these drugs would be expected to have an effect. The classification results at the individual level

indicate that the PPS demarcation formula could also be useful in clinical practice, as a companion diagnostic and a step towards true personalized medicine.

### Acknowledgements

A.T., K.V.R., J.B., and J.G.P. conceived and designed the project. K.V.R., J.B., J.G., and H.P.F.K. performed the experiment. F.M., J.V.H., P.J., L.D.L., J.-J.M., W.E., H.W.F., and B.O. supervised the experiment. A.T., K.B.E.B., D.H., R.P.J.D.L., K.V.R., and R.K. analysed the data. A.T. and J.G.P. wrote the initial draft of the manuscript, with all other authors contributing to editing into the final form.

### Declaration of conflicting interests

A.T. is the CEO of Emotional Brain (EB) and a shareholder of EB. F.M. and P.J. are consultants to EB. K.B.E.B. and D.H. are advisors to EB. J.V.H. reports no conflict of interest. R.P.J.D.L., J.-J.M., and W.E. are advisors to EB and own shares/stock options in EB. K.V.R., J.B., J.G., and H.P.F.K. are employees of EB and own shares/stock options in EB. R.K. is an employee of EB. L.D.L. is a consultant to EB and owns shares/stock options in EB. B.O. is a member of the scientific advisory board of EB and owns shares/stock options in EB. H.W.F. is an advisor to EB, and his employer has a licence agreement with EB. J.G.P. is on the advisory board of, and/or a consultant to, EB, Palatin Technologies, and Acadia Pharmaceuticals and has received research operating grants from the Canadian Institutes for Health Research and the Natural Sciences and Engineering Research Council of Canada.

### Funding

This study was funded by Emotional Brain BV, Almere, The Netherlands.

### Supplementary material

Supplementary material is available for this article online.

### References

- Mitchell KR, Mercer CH, Ploubidis GB, et al. Sexual function in Britain: findings from the third national survey of sexual attitudes and lifestyles (Natsal-3). *Lancet* 2013; 382: 1817–1829.
- Shifren JL, Monz BU, Russo PA, et al. Sexual problems and distress in United States women: prevalence and correlates. *Obstet Gynecol* 2008; 112: 970–978.
- Davison SL, Bell RJ, LaChina M, et al. The relationship between self-reported sexual satisfaction and general well-being in women. *J Sex Med* 2009; 6–2697: 2690.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Association, 2013.
- Pfaus JG. Pathways of sexual desire. *J Sex Med* 2009; 6: 1506–1533.
- Kingsberg SA, Clayton AH and Pfaus JG. The female sexual response: current models, neurobiological underpinnings and agents currently approved or under investigation for the treatment of hypoactive sexual desire disorder. *CNS Drugs* 2015; 29: 915–933.
- Brotto L, Atallah S, Johnson-Agbakwu C, et al. Psychological and interpersonal dimensions of sexual function and dysfunction. *J Sex Med* 2016; 13: 538–571.
- Hayes RD, Dennerstein L, Bennett CM, et al. Relationship between hypoactive sexual desire disorder and aging. *Fertil Steril* 2007; 87: 107–112.
- West SL, D'Aloisio AA, Agans RP, et al. Prevalence of low sexual desire and hypoactive sexual desire disorder in a nationally representative sample of US women. *Arch Intern Med* 2008; 168: 1441–1449.
- Bruzziches R, Francomano D, Gareri P, et al. An update on pharmacological treatment of erectile dysfunction with phosphodiesterase type 5 inhibitors. *Expert Opin Pharmacother* 2013; 14: 1333–1344.
- Joffe HV, Chang C, Sewell C, et al. FDA approval of flibanserin – treating hypoactive sexual desire disorder. *N Engl J Med* 2016; 374: 101–104.
- Bloemers J, van Rooij K, Poels S, et al. Toward personalized sexual medicine (Part 1): integrating the “Dual Control Model” into differential drug treatments for hypoactive sexual desire disorder and female sexual arousal disorder. *J Sex Med* 2013; 10: 791–809.
- Poels S, Bloemers J, van Rooij K, et al. Toward personalized sexual medicine (part 2): testosterone combined with a PDE5 inhibitor increases sexual satisfaction in women with HSDD and FSAD, and a low sensitive system for sexual cues. *J Sex Med* 2013; 10: 810–823.
- van Rooij K, Poels S, Bloemers J, et al. Toward personalized sexual medicine (part 3): testosterone combined with a serotonin1A receptor agonist increases sexual satisfaction in women with HSDD and FSAD, and dysfunctional activation of sexual inhibitory mechanisms. *J Sex Med* 2013; 10: 824–837.
- van der Made F, Bloemers J, Yassem WE, et al. The influence of testosterone combined with a PDE5-inhibitor on cognitive, affective, and physiological sexual functioning in women suffering from sexual dysfunction. *J Sex Med* 2009; 6: 777–790.
- Bancroft J and Janssen E. The dual control model of male sexual response: a theoretical approach to centrally mediated erectile dysfunction. *Neurosci Biobehav Rev* 2000; 24: 571–579.
- van der Made F, Bloemers J, van Ham D, et al. Childhood sexual abuse, selective attention for sexual cues and the effects of testosterone with or without vardenafil on physiological sexual arousal in women with sexual dysfunction: a pilot study. *J Sex Med* 2009; 6: 429–439.
- Gerritsen J, van der Made F, Bloemers J, et al. The clitoral photoplethysmograph: a new way of assessing genital arousal in women. *J Sex Med* 2009; 6: 1678–1687.
- Tuiten A, van Rooij K, Bloemers J, et al. Efficacy and safety of on-demand use of 2 treatments designed for different etiologies of female sexual interest/arousal disorder: 3 randomized clinical trials. *J Sex Med* 2018; 15: 201–216.
- Bloemers J, Scholte HS, van Rooij K, et al. Reduced gray matter volume and increased white matter fractional anisotropy in women with hypoactive sexual desire disorder. *J Sex Med* 2014; 11: 753–767.

21. Tuiten A, Van Honk J, Koppeschaar H, et al. Time course of effects of testosterone administration on sexual arousal in women. *Arch Gen Psychiatry* 2000; 57: 149–153, discussion 155–156.
22. Tuiten A, van Honk J, Verbaten R, et al. Can sublingual testosterone increase subjective and physiological measures of laboratory-induced sexual arousal? *Arch Gen Psychiatry* 2002; 59: 465–466.
23. Purcell SM, Wray NR, Stone JL, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; 460: 748–752.
24. Yang J, Manolio TA, Pasquale LR, et al. Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet* 2011; 43: 519–525.
25. Visscher PM, Goddard ME, Derks EM, et al. Evidence-based psychiatric genetics, AKA the false dichotomy between common and rare variant hypotheses. *Mol Psychiatry* 2012; 17: 474–485.
26. Gunst A, Jern P, Westberg L, et al. A study of possible associations between single nucleotide polymorphisms in the estrogen receptor 2 gene and female sexual desire. *J Sex Med* 2015; 12: 676–684.
27. Burri A, Hysi P, Clop A, et al. A genome-wide association study of female sexual dysfunction. *PLoS ONE* 2012; 7: e35041.
28. Visscher PM, Hill WG and Wray NR. Heritability in the genomics era – concepts and misconceptions. *Nat Rev Genet* 2008; 9: 255–266.
29. Gibson G. Rare and common variants: twenty arguments. *Nat Rev Genet* 2012; 13: 135–145.
30. Fisher RA. The correlation between relatives on the supposition of Mendelian inheritance. *Trans R Soc Edinb* 1918; 52: 399–433.
31. Fisher RA. 1930 *The genetical theory of natural selection*. Oxford: The Clarendon Press.
32. Fisher RA. Adaptation and mutations: a lecture to the science masters' association. *Sch Sci Rev* 1934; 15: 294–301.
33. Plomin R, Haworth CMA and Davis OSP. Common disorders are quantitative traits. *Nat Rev Genet* 2009; 10: 872–878.
34. Lim S-W, Won H-H, Kim H, et al. Genetic prediction of antidepressant drug response and nonresponse in Korean patients. *PLoS ONE* 2014; 9: e107098.
35. Eisenegger C, Knoch D, Ebstein RP, et al. Dopamine receptor D4 polymorphism predicts the effect of L-DOPA on gambling behavior. *Biol Psychiatry* 2010; 67: 702–706.
36. Eisenegger C, Naef M, Linssen A, et al. Role of dopamine D2 receptors in human reinforcement learning. *Neuropsychopharmacology* 2014; 39: 2366–2375.
37. van Nes Y, Bloemers J, van der Heijden PGM, et al. The Sexual Event Diary (SED): development and validation of a standardized questionnaire for assessing female sexual functioning during discrete sexual events. *J Sex Med* 2017; 14: 1438–1450.
38. van Nes Y, Bloemers J, Kessels R, et al. Psychometric properties of the sexual event diary in a sample of Dutch women with female sexual interest/arousal disorder. *J Sex Med* 2018; 15: 722–731.
39. Bloemers J, van Rooij K, de Leede L, et al. Single dose sublingual testosterone and oral sildenafil vs. a dual route/dual release fixed dose combination tablet: a pharmacokinetic comparison: pharmacokinetics of testosterone and sildenafil combination tablet. *Br J Clin Pharmacol* 2016; 81: 1091–1102.
40. van Rooij K, de Leede L, Frijlink HW, et al. Pharmacokinetics of a prototype formulation of sublingual testosterone and a buspirone tablet, versus an advanced combination tablet of testosterone and buspirone in healthy premenopausal women. *Drugs RD* 2014; 14: 125–132.
41. Dunlop W, Cortina J, Vaslow J, et al. Meta-analysis of experiments with matched groups or repeated measures designs. *Psychol Methods* 1996; 1: 170–177.
42. Kraemer HC and Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. *Biol Psychiatry* 2006; 59: 990–996.