

The design and methodology of premature ejaculation interventional studies

Chris G. McMahon

Australian Centre for Sexual Health, Sydney, NSW 2065, Australia

Correspondence to: Chris G. McMahon. Australian Centre for Sexual Health, Sydney, NSW 2065, Australia. Email: chrismcmahon@gmail.com.

Abstract: Large well-designed clinical efficacy and safety randomized clinical trials (RCTs) are required to achieve regulatory approval of new drug treatments. The objective of this article is to make recommendations for the criteria for defining and selecting the clinical trial study population, design and efficacy outcomes measures which comprise ideal premature ejaculation (PE) interventional trial methodology. Data on clinical trial design, epidemiology, definitions, dimensions and psychological impact of PE was reviewed, critiqued and incorporated into a series of recommendations for standardisation of PE clinical trial design, outcome measures and reporting using the principles of evidence based medicine. Data from PE interventional studies are only reliable, interpretable and capable of being generalised to patients with PE, when study populations are defined by the International Society for Sexual Medicine (ISSM) multivariate definition of PE. PE intervention trials should employ a double-blind RCT methodology and include placebo control, active standard drug control, and/or dose comparison trials. Ejaculatory latency time (ELT) and subject/partner outcome measures of control, personal/partner/relationship distress and other study-specific outcome measures should be used as outcome measures. There is currently no published literature which identifies a clinically significant threshold response to intervention. The ISSM definition of PE reflects the contemporary understanding of PE and represents the state-of-the-art multi-dimensional definition of PE and is recommended as the basis of diagnosis of PE for all PE clinical trials.

Keywords: Treatment; premature ejaculation (PE); trial methodology

Submitted Jan 31, 2016. Accepted for publication Mar 07, 2016.

doi: 10.21037/tau.2016.03.28

View this article at: <http://dx.doi.org/10.21037/tau.2016.03.28>

In the last 2 decades an increasing number of publications have reported the pharmacological treatment of premature ejaculation (PE) with a variety of different medications which act either centrally or locally to retard the psychoneurological control of ejaculation and subsequent orgasm. The introduction of SSRIs in the early 1990s revolutionised the treatment of PE (1). Multiple well-controlled evidence-based studies have demonstrated the efficacy and safety of daily or on-demand administration of SSRIs in delaying ejaculation, confirming their role as first-line agents for the treatment of lifelong and acquired PE (2-4). The pharmaceutical industry has finally developed an interest in the identification of potential therapeutic targets and the development of PE pharmacotherapy. The PE treatment paradigm, previously limited to behavioural

psychotherapy, has progressively expanded to include drug treatment (5,6).

Regulatory approval, ethical and responsible marketing by the pharmaceutical industry, and confident, efficacious and safe prescribing by the treating physician of a new drug treatment for PE demand initial evaluation of investigational drugs in large industry-funded clinical efficacy and safety phase II, III trials and subsequent phase IV post-regulatory approval industry-funded or independent pharmaco-vigilance trials. However, data from PE observational, interventional and treatment preference studies are only reliable, interpretable and capable of being generalised to patients with PE, when study populations are rigidly defined using evidence based multivariate definition of PE, and intervention outcomes are measured with

consistent objective physiological measures or sensitive, validated outcome assessment instruments (7).

The article reviews and critiques data on clinical trial design, epidemiology, definitions, dimensions and psychological impact of PE to make a series of recommendations for the criteria for defining and selecting the clinical trial study population, design and efficacy outcomes measures which comprise ideal PE intervention trial methodology.

Intervention studies

In interventional studies, the effect of a particular drug or other intervention upon the health of the research subjects is measured. Pre-clinical studies, comprising *in-vitro* and animal studies, and phase 0 trials, exploratory, first-in-human usually PK studies to expedite the drug by establishing very early on whether the drug or agent behaves in human subjects as was anticipated from preclinical studies, will not be considered in this article. Phase I trials assess safety, tolerability, pharmacokinetics (PK) and pharmacodynamics (PD) of an investigational drug with dose-escalation in small groups of usually healthy volunteers. Phase II trials investigate efficacy and dosing requirements in randomized clinical trials (RCTs) of larger groups of subjects [20–300]. Phase III trials explore efficacy, safety and drug–drug interactions in general and special populations in large RCTs. Phase IV trials involve post-marketing pharmacovigilance and ongoing technical support and are often requested by regulatory agencies.

Ethical human experimentation guidelines

PE intervention clinical trials must be conducted with strict adherence to the International Committee of Harmonization (ICH) Good Clinical Practice (GCP) Guidelines (8,9). The Declaration of Helsinki was developed by the World Medical Association (WMA), as a set of ethical principles for the medical community regarding human experimentation (10). It is widely regarded as the cornerstone document of human research ethics although it is not a legally binding instrument in international law. It draws its authority from the degree to which it has been codified in, or influenced national or regional legislation and regulations. Key Aspects of ICH GCP guidelines include approval of the trial protocol by a properly constituted institutional review board (IRB) or an independent ethics committee (IEC) and a subject informed

consent which may include a separate consent for non-mandatory pharmacogenomic testing and a subject's partner informed consent.

Clinical trial methodology

The basis of ideal PE clinical trial design involves adequately defining the trial population, a double-blind placebo-controlled interventional RCT methodology, and the use of the sensitive, validated and reproducible outcome measures.

Determination of trial sample size

The number of subjects or the trial or sample size must be sufficient to maximize the probability of the trial results achieving statistical significance and minimize the incidence of false-positives (alpha, or type 1, errors) and false-negatives errors (beta, or type 2, errors). This is often referred to as “powering the trial” and is reliant upon a series of assumptions based upon prior experience with the disease or interventions being compared to make estimates of the likely effects that can be expected.

In a superiority trial, significance testing focuses on asserting the null hypothesis, the hypothesis that there is no true difference among the compared groups. Similarly, a non-inferiority design is one in which the assertion is that a trial is unable to detect a superior or equivalent result. In these circumstances, if no difference is found among the trial groups, it may be because there is either no true difference or the trial was not sufficiently large, and an outcome indicating similarity occurred purely by chance.

To decide whether there is no difference or whether the similarity was by chance, the alpha level is established, representing the maximum probability of making a false-positive error that is acceptable. In general, the alpha level is set at $P=0.05$, so there is no more than a 1 in 20 probability that the outcome has occurred by chance. However, when interim analyses have been performed or when multiple comparisons are made in data, a higher and therefore more rigorous alpha level should be sought. This adjustment for multiple comparisons, known as the Bonferroni correction, aims to limit the possibility that if 20 statistical comparisons are made with an error probability of 1 in 20, then by chance 1 will be an erroneous false-positive result.

When mean differences are compared, such as with a *t*-test, and a *P* value of less than 0.05 is observed, there is little interest in the false-negative, or beta level. However, if

the sample size is too small and a non-significant alpha level is obtained, it may have been caused by a false-negative result. By convention, when designing trials, the beta level, the acceptable level for getting a false-negative result, is set at 20%, i.e., the trial will have a 20% chance of missing a true-positive finding. The smaller the beta level the investigator is willing to accept, the larger the sample size needed for the trial to be adequately powered.

Defining the trial population

The trial population must be representative of patients with PE and their partners and is defined by a series of inclusion and exclusion criteria (*Table 1*). The trial population should either be limited to subjects with the same sexual orientation and partner gender preference or subjects with differing sexual orientation and partner gender preference should be treated as separate trial populations.

Classification of premature ejaculation (PE)

The population of men with PE is heterogeneous and varies in patient demographics, symptoms, risk factors and pathophysiology. In 1943, Shapiro (12) proposed classification of PE into two types, types B and A. In 1989, In 1989 Godpodinoff (13) renamed both types as lifelong (primary) and acquired (secondary) PE. Over the years, other attempts to specify subtypes have occurred (e.g., global *vs.* situational, the effect of a substance, etc.).

Lifelong PE is a syndrome characterized by a cluster of core symptoms including early ejaculation at nearly every intercourse within 30–60 s in the majority of cases (80%) or between 1–2 min (20%), with every or nearly every sexual partner and from the first sexual encounters onwards (14,15).

Acquired PE differs in that sufferers develop early ejaculation at some point in their life, which is often situational, having previously had normal ejaculation experiences. The main distinguishing features between presentations of these two syndromes are the time of onset of symptoms and the reduction in previously normal ejaculatory latency of acquired PE.

Community based normative intravaginal ejaculatory latency time (IELT) research and observational studies of men with PE demonstrated that although IELTs of less than 1 min have a low prevalence of about 2.5% in the general population, a substantially higher percentage of men with normal IELT complain of PE (16–18). In order to take account of this diversity, Waldinger and Schweitzer

(19,20) proposed a new classification of PE in which four PE subtypes are distinguished on the basis of the duration of the IELT, frequency of complaints, and course in life. In addition to lifelong PE and acquired PE, this classification includes natural variable PE (or variable PE) and premature-like ejaculatory dysfunction (or subjective PE). Men with variable PE occasionally experience an early ejaculation. It should not be regarded as a disorder, but as a natural variation of the ejaculation time in men (21). On the other hand, men with subjective PE complain of PE, while actually having a normal or even extended ejaculation time (21). The complaint of PE in these men is probably related to psychological and/or cultural factors. In contrast, the consistent early ejaculations of lifelong PE suggested an underlying neurobiological functional disturbance, whereas the early ejaculation of acquired PE is more related to underlying medical causes. Serefoglu *et al.* (22,23) confirmed the existence of these four PE subtypes in a cohort of men in Turkey. Recently, Zhang *et al.* (24) and Gao *et al.* (25) using a similar methodology confirmed similar prevalence rates of the four PE subtypes in China to that reported by Serefoglu *et al.* (22,23). This new classification and continued research into the diverse phenomenology, etiology and pathogenesis of PE is expected to provide a better understanding of the four PE subtypes (19).

The erroneous assumption that PE is the most common male sexual disorder, and the variance between the incidence of PE in observational studies which rely on self-diagnosis or DSM-IV-TR diagnosis, and that reported in community stopwatch IELT studies most likely reflects the community incidence of natural variable PE and/or premature life ejaculatory dysfunction rather than the actual syndromes of lifelong and acquired PE (18,26,27). As such, different PE sub-types should be treated as demographically and etiologically distinct disorders and analysed as separate PE subgroups in clinical trials. Drug intervention studies should be limited to subjects with either lifelong PE or etiology specific acquired, e.g., PE with comorbid ED, comorbid prostatitis etc.

Definitions of premature ejaculation (PE)

Research into the treatment and epidemiology of PE is heavily dependent on how PE is defined. Prior to 2007, the medical literature contains several univariate and multivariate operational definitions of PE (6,28–35). Each of these definitions characterise men with PE using all

Table 1 Inclusion and exclusion criteria for PE intervention trial

Inclusion criteria

Male subjects aged ≥ 18 years

Subject has provided written informed consent

Subjects must be in a stable, monogamous, sexual relationship with the same partner for at least 6 months and plan to maintain this relationship for the duration of the trial. Studies should either be limited to subjects with the same sexual orientation and partner gender preference, or treat subjects with same sex partners as a separate trial group. Implicit in trial design is the recognition that trial endpoints may be specific to the subjects preferred method of sexual expression, e.g., anal intercourse, felatio, solitary or mutual masturbation etc.

Subjects must be in good general health with no clinically relevant abnormalities, normal blood chemistry, testosterone, CBC, urinalysis, and 12-lead ECG

Subjects must be prepared to attempt intercourse on a regular basis and at least once a week

Partners must have a negative urine pregnancy test

Subjects must meet criteria for diagnosis of PE using a multivariate definition of PE and a baseline trial specific threshold ejaculatory latency time, e.g., intravaginal ejaculatory latency time (IELT), intra-anal ejaculatory latency time (AELT), intra-oral ejaculatory latency time (OELT) or masturbatory ejaculatory latency time (MELT)

Subjects must meet criteria for diagnosis of PE with a PEDT score ≥ 11 indicates

Exclusion criteria

Subject must not have used any other investigational drug within the past 1 month or within a period of less than 5 times the drug's half-life, whichever is longer

Subject must not have a significant history of current cardiovascular, pulmonary, gastrointestinal, hematologic, neurologic, locomotor, immunologic, ophthalmologic, metabolic, endocrine, thromboembolic, rheumatologic, oncologic, renal, or hepatic disorders

Subject must not have a history pelvic/retroperitoneal surgery or radiotherapy, multiple sclerosis, spinal cord injury or prostatitis which may be associated with the onset of PE symptoms and considered a potential cause of PE

Subject must not be seropositive for the human immunodeficiency virus (HIV), hepatitis B surface antigen (HBsAg), or hepatitis C

Subject must not have a current or past history of depressive or anxiety disorder, dysthymia, suicidality, (hypo) manic episode, panic disorder, agoraphobia, social phobia, obsessive-compulsive disorder, posttraumatic stress disorder, or psychotic disorders. This should be determined using a validated mental health screening inventory e.g., Mini International Neuropsychiatric Interview (M.I.N.I.) (11)

Subject with a current or past history of alcohol abuse and dependence, non-alcohol psychoactive substance use disorder

Subject prescribed any medication which may cause sexual dysfunction as an adverse effect and/or is contraindicated or has a reported interaction with the trial drug.

Subject with hypoactive sexual desire, retrograde, delayed or absent orgasm or ejaculation or ED (IIEF EF domain < 26 , IIEF-5 > 21)

Subject with hypogonadism, hyperprolactinemia or untreated or insufficiently treated hypothyroidism

Partner of subject with clinically significant female/male sexual dysfunction (FSD/MSD) including hypoactive sexual desire and dyspareunia which may significantly impact the sexual relationship with the subject

or most of the accepted dimensions of this condition: ejaculatory latency, perceived ability to control ejaculation, and negative psychological consequences of PE including reduced sexual satisfaction, personal distress, partner distress and interpersonal or relationship distress. The major criticisms of the extant definitions included their failure to be evidenced-based, lack of specific operational criteria, excessive vagueness, and reliance on the subjective

judgment of the diagnostician.

This lack of agreement as to what constitutes PE hampered clinical research into the etiology and management of this condition. This potential for errors in the diagnosis of PE was demonstrated in two observational studies in which PE was diagnosed solely by the application of the DSM-IV-TR definition (18,36). Giuliano *et al.* (36) diagnosed PE using DSM-IV-TR criteria in 201 of 1,115

subjects (18%) and predictably reported that the mean and median IELT was lower in subjects diagnosed with PE compared to non-PE subjects. There is, however, substantial overlap in stop-watch IELT values between the two groups. The authors acknowledge this and interpret this as demonstrating that IELT as a single measure cannot discriminate subjects with from subjects without PE. However, this overlap most likely relates to the use of the flawed DSM-IV-TR definition and represents a substantial tautological error in trial design. In subjects diagnosed with PE, the IELT range extended to almost 28 min at 8 weeks with 44.3% of subjects having an IELT in excess of 2 min and 24.9% of subjects exceeding 4 min. In the non-PE group, 12.1% of subjects had an IELT less than 2 min and a small number of subjects had an IELT of 0, suggesting antiportal ejaculation and therefore severe PE. As such, this trial demonstrates that a subject diagnosed as having PE according to the DSM-IV-TR criteria has a 44.3% risk of not having PE if a PE diagnostic threshold IELT of 2 min is used. If we are to assume that the conclusions of the community-based normative IELT trial are correct, these observational trials must be regarded as primarily measures of the reliability of DSM-IV-TR as a diagnostic tool for PE (17). Furthermore, conclusions regarding the relationship between patient-reported outcomes (PROs) and IELT based on data from this inadequately selected trial group must be regarded with some caution as they are more assumptive than evidence-based and cannot be reliably generalized to subjects with this condition.

The inadequate methodology of these studies related to the use of DSM-IV-TR as a diagnostic tool and the impact of that methodology upon the author's conclusions are testament to the importance of a multidimensional evidence-based definition of PE.

International Society of Sexual Medicine (ISSM) definition of premature ejaculation (PE)

In the last decade, substantial progress has been made in the development of evidence-based methodology for PE epidemiologic and drug treatment research using the objective IELT and subjective validated PRO measures. Evidence-based definitions seek to limit errors of classification and thereby increase the likelihood that existing and newly developed therapeutic strategies are truly effective in carefully selected dysfunctional populations (30).

ISSM developed the first contemporary, evidence-based definition of lifelong PE in 2007 (37). In 2013, ISSM developed a contemporary, evidence-based definition single

unifying definition of both acquired and lifelong PE (38). Members unanimously agreed that although lifelong and acquired PE are distinct and different demographic and etiological populations, they can be jointly defined, in part, by the constructs of time from penetration to ejaculation, inability to delay ejaculation and negative personal consequences from PE. The committee agreed that the presence of these mutual constructs was sufficient justification for the development of a single unifying definition of both lifelong and acquired PE. Finally, the committee determined that the presence of a clinically significant and bothersome reduction in latency time, often to about 3 min or less was an additional key defining dimension of acquired PE.

The second Ad Hoc ISSM Committee for the Definition of Premature Ejaculation [2013] defined premature ejaculation (lifelong and acquired PE) as a male sexual dysfunction characterized by:

- (I) Ejaculation which always or nearly always occurs prior to or within about 1 min of vaginal penetration (lifelong PE), or, a clinically significant and bothersome reduction in latency time, often to about 3 min or less (acquired PE);
- (II) The inability to delay ejaculation on all or nearly all vaginal penetrations;
- (III) Negative personal consequences, such as distress, bother, frustration and/or the avoidance of sexual intimacy.

The unified ISSM definition of lifelong and acquired PE represents the first evidence-based definition for these conditions. This definition should form the basis for the office diagnosis of lifelong PE and the design of PE observational and interventional clinical trials. It is limited to men engaging in vaginal intercourse as there are few studies available on PE research in homosexual men or during other forms of sexual expression. This definition intentionally includes a degree of diagnostic conservatism and flexibility. The 1-minute IELT cut-off point for lifelong PE should not be applied in the most absolute sense, as about 10% of men seeking treatment for lifelong PE have IELTs of 1–2 min. The phrase, "within about 1 minute" must be interpreted as giving the clinician sufficient flexibility to diagnose PE also in men who report an IELT as long as 90 s. Similarly, a degree of flexible clinical judgement is key to the recognition and interpretation of a bothersome change in ejaculatory latency with reduction of pre-morbid latency to ≤ 3 min in men with acquired PE. Men who report these ejaculatory latencies

Table 2 Basic design of premature ejaculation (PE) intervention trials

Trial methodology	Interventional trial	
	Drug	Psychotherapeutic
Trial population	<ol style="list-style-type: none"> 1. Lifelong PE; 2. Etiology specific acquired PE 	<ol style="list-style-type: none"> 1. Psychogenic acquired PE; 2. Natural variable PE; 3. PE-like ejaculatory dysfunction (EjD)
Inclusion/exclusion criteria	<ol style="list-style-type: none"> 1. Intravaginal ejaculatory latency time (IELT) ≤ 1 min; 2. International Society of Sexual Medicine (ISSM) PE definition; 3. PE diagnostic inventory; 4. IIEF EF ≥ 26 or ≤ 26 for acquired PE with ED; 5. See <i>Table 1</i>-Part 1* 	<ol style="list-style-type: none"> 1. IELT ≤ 1 min; 2. ISSM PE definition; 3. PE diagnostic inventory; 4. IIEF EF ≥ 26 or ≤ 26 for acquired PE with ED; 5. See <i>Table 1</i>-Part 1*
Trial design	<ol style="list-style-type: none"> 1. Double-blind, parallel-arm placebo-controlled 12-week; 2. Head-to-head comparator cross-over placebo-controlled 12-week 	<ol style="list-style-type: none"> 1. Psychotherapy control; 2. Trial specific
Outcome measures	<ol style="list-style-type: none"> 1. Stopwatch IELT; 2. Single item PROs for control, satisfaction, personal/interpersonal distress; 3. Multi-domain PE inventory; 4. Partner reported outcomes (PaROs); 5. Trial specific PROs; 6. Adverse events; 7. Discontinuation rate 	<ol style="list-style-type: none"> 1. Stopwatch IELT; 2. Single item patient-reported outcomes (PROs) for control, satisfaction, personal/interpersonal distress; 3. Multi-domain PE inventory; 4. Partner reported outcomes (PaROs); 5. Trial specific PROs; 6. Discontinuation rate

*, the inclusion/exclusion criteria listed in *Table 1*.

but describe adequate control and no personal negative consequences related to their rapid ejaculation do not merit the diagnosis of PE.

Diagnostic and statistical manual of mental disorders (DSM-5) definition of premature ejaculation (PE)

Based upon the same data that supported the ISSM definition of lifelong PE, the recently published DSM-5 definition of PE (39) now includes an objective ejaculatory latency criterion. DSM-5 defines PE as “...a persistent or recurrent pattern of ejaculation occurring during partnered sexual activity within approximately 1 minute following vaginal penetration and before the individual wishes it. This symptom must have been present for at least 6 months and must be experienced on almost all or all (approximately 75–100%) occasions of sexual activity. It causes clinically significant distress in the individual.” (39). The DSM-5 definition of

PE requires clinicians to specify PE as either lifelong or acquired, and as generalized or situational. In addition, the DSM-5 definition of PE distinguishes between mild PE (ejaculation occurring within approximately 30 s to 1 min of vaginal penetration), moderate PE (ejaculation occurring within approximately 15–30 s of vaginal penetration) and severe PE (ejaculation occurring prior to sexual activity, at the start of sexual activity, or within approximately 15 s of vaginal penetration).

Trial design

Intervention studies comprise drug intervention studies, psychotherapeutic intervention studies or studies combining both interventions. The basic elements of ideal PE intervention study are summarized in *Table 2*.

Ideal well-controlled PE RCT methodology includes

placebo control, active standard drug control, and dose comparison trials. Historical control and unblinded no treatment concurrent control are not valid in PE interventional trials. The goal of placebo control and dose comparison trials is to show superiority of the new drug, at some dose, over either placebo or a lower dose of that same drug. The goal of an active standard drug controlled trial is to show either the superiority or equivalence, i.e., non-inferiority of an investigational drug to an active standard. Historically, industry-funded trials avoid active standard drug controlled trial methodology when they are unlikely to reveal the superiority of a new drug over available treatments (40). In general, any of the superiority designs, whether the comparison is to placebo, to a lower dose of the same drug, or an active standard drug, is considered scientifically valid.

The most important reason for using a placebo control in a RCT is to act as an internal validation. The response to an investigational drug is evaluated by comparison to the response to placebo. The difference in responses between the investigational drug group and the placebo group reflects the activity of the investigational drug. This assumes that the randomisation process has ensured that all other factors that might potentially affect response are equally distributed between the groups. In general, the placebo response is lower and more stable in single centre studies and higher and more variable in multicenter studies which probably reflect the contribution of different confounding factors in different centres. A randomised comparison between investigational drug and placebo is particularly necessary if the therapeutic effect of a treatment is small or variable and the response to placebo is high or differs in different settings. Placebo-controlled studies provide the most unequivocal evidence of efficacy. Evidence from at least two positive well designed and conducted placebo-controlled studies is generally accepted as appropriate to establish the efficacy of a drug.

However, the nature of placebo is incompletely understood and its use as a control cannot be based upon the assumption that it is an ineffective inert substance. Subjects who receive a placebo are provided with support and concern and are reassured by the perception that the complaint is understood and is being taken seriously. This serves to encourage a positive subject attitude that strengthens the therapeutic alliance and may constitute a de-facto form of cognitive behaviour therapy. The placebo response encompasses a variety of these non-specific factors and, will, in subjects receptive to reassurance, combine to produce clinical improvement in subjects with

sexual dysfunction with any degree of psychogenic causality which is usually time-limited.

A fundamental aspect of trial design is the inverse relationship between trial sample size and the population effect size that is detectable. If an investigational drug is initially evaluated in a placebo controlled trial, the number of subjects that must be enrolled is relatively small and, consequently, the number of non-responders is also small. In contrast, because there would be small differences between an investigational drug and another active medication, more subjects must be enrolled and paradoxically there would be more non-responders in a RCT that does not include placebo controls.

Consider the sample size necessary to detect population treatment differences with a statistical power of 0.80 and an alpha level set at $P=0.05$ in two different types of trial design. In a placebo controlled trial, the sample size necessary to detect a difference between a 30% placebo response rate and a 60% investigational drug response rate is 48 subjects per group. In contrast, when 2 active medications are compared where the difference in response rates is likely to be smaller, 407 subjects per group would be required to detect a difference between a 60% investigational drug response rate and 50% standard drug response rate. For this reason, a direct comparison with an effective standard therapy should not be conducted unless the investigational drug has been shown to be superior to placebo.

Drug intervention trials in PE should control for psychotherapy, physician-patient alliance, and inter-current life events, since these variables may have significant potential to influence clinical outcome, and currently receive little attention in trial methodology (41-43). The assumption that psychotherapy does not directly confound the effects of pharmacologic intervention in PE is inconsistent with knowledge about direct effects of psychotherapy on brain physiology, and other findings such as the influence of psychosocial factors on gene expression (44-46). The recent proposal that lifelong PE has a neurobiological basis has led the field to an unsubstantiated assumption that biological phenomena are more relevant in our search for the truth than psychotherapy (47). Clearly, control for psychotherapy is often challenging as psychotherapy is more difficult to standardize and quantify than demographic or biological clinical variables but this consideration does not apply to exclusion of all concomitant psychotherapy. Psychotherapy prior to drug trial enrolment or uncontrolled psychotherapy during the treatment phase of a trial must be regarded as significant mediators of

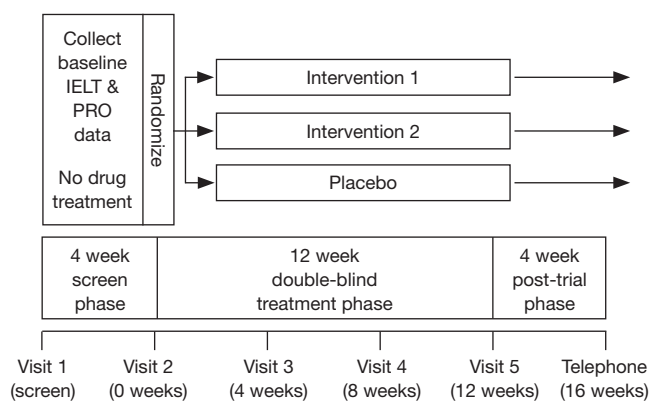


Figure 1 Proposed trial flow chart for a PE intervention trial.

outcome (43). Lack of control of this confounding variable could compromise internal validity if subjects in one group or the other happen, despite randomization, to receive more or better psychotherapy. It is likely that uncontrolled receipt of psychotherapy by subjects results in treatment effects that dilute differences between active and control pharmacotherapy treatments. Furthermore, uncontrolled psychotherapy may also confound with pharmacotherapy assignment, since subjects randomized to placebo or a less effective treatment may be more likely to seek adjunctive psychotherapy. This would further dilute differences between active drug and control in placebo-controlled studies. Adequate control for these potential confounding factors pharmacotherapy controlled trial may lead to a more valid determination of treatment outcome in PE trials.

Phases of premature ejaculation (PE) interventional studies

Interventional studies must comprise 3 distinct phases: an initial treatment free screening phase, a treatment phase and a post-trial phase (*Figure 1*).

Screening phase

Subjects must provide written informed consent and satisfy the study specific inclusion and exclusion criteria (*Table 1*). Inclusion criteria should include compliance with a multivariate definition of PE. The diagnosis of PE using the premature ejaculation diagnostic tool (PEDT) has less discriminating power and physician diagnosis alone has no role in PE intervention trial design (38,39,48). During the initial treatment free screening phase of approximately 4 weeks, subject and partner demographic data, baseline ELT, erectile and overall sexual function, PROs and partner

reported outcomes (PaROs) and other trial specific outcome measures should be collected in order to determine whether inclusion/exclusion criteria are satisfied and to establish candidacy for enrolment in the trial. Subjects should be assigned a randomization number either manually or via a central interactive voice response system (IVRS), which is derived from a randomization table generated by the method of random permuted blocks and assigned to a treatment arm (49).

Premature ejaculation diagnostic tool (PEDT)

The PEDT was developed specifically for use as a screening questionnaire (48,50,51). This questionnaire is a brief, 5-item 0–25 score questionnaire used to screen men for potential presence of PE based on DSM-IV-TR criteria of lack of control, frequency, minimal stimulation, distress and interpersonal difficulty. It has good convergent validity and re-test reliability and sensitivity/specificity analysis suggests that a score ≥ 11 indicates PE (50). The PEDT is limited in several respects, but represents a significant development towards simplifying the methodology of PE drug studies. Several authors have confirmed the validity of the PEDT as a diagnostic tool in the clinical setting (52,53). Future development of the PEDT as diagnostic tool would be incomplete without further validation of this tool to determine the potential relationship between score, severity of PE and response to treatment. The PEDT is not validated as an intervention outcome measure.

Assessment of baseline ejaculatory latency time (ELT)

Measurement of the ELT by stopwatch is the best method to measure treatment response and should be used as a primary efficacy outcome measure. ELT is the length of time between vaginal (IELT), anal ejaculatory latency time (AELT), oral penetration ejaculatory latency time (OELT) or the initiation of masturbation ejaculatory latency time (MELT), and forms the basis of determining candidacy for enrolment in PE interventional clinical trials and should be used as a primary efficacy endpoint (2).

The IELT, AELT, OELT and MELT can be estimated by the subject or measured with a stopwatch operated by the female partner, is expressed in seconds or minutes and in cases of ante-portal ejaculation is equal to zero. Several authors report that estimated and stopwatch IELT correlate reasonably well or are interchangeable in assigning PE status when estimated IELT is combined with PROs (54–56). Whilst these findings do provide support for the use of self-estimation of IELT for the diagnosis of PE in clinical practice, it has a limited role in intervention clinical

trials but fails to provide sufficient objective evidence for determination and comparison of treatment outcome in interventional studies. As such, stopwatch IELT forms the basis of diagnosing PE in interventional clinical trials (2,33).

The ISSM definition of PE nominates a threshold stopwatch IELT cutoff point for diagnosis of lifelong PE of “...prior to or within about one minute of vaginal penetration” and “...a clinically significant and bothersome reduction in latency time, often to about 3 minutes or less” for acquired PE (38).

Baseline IELT should be determined during a 4-week baseline period during which the subject should have at least 4 intercourses at least 24 h apart. Determination of IELT should be limited to the first intercourse attempt and not to subsequent attempts during the same period of sexual contact. Use of condoms, topical anaesthetic creams or prior significant consumption of alcohol is not permitted.

As the IELT in both the general population and in subjects with PE is distributed in a positively skewed pattern, reporting baseline and trial-end IELTs as arithmetic means over-estimates the treatment response by as much as 45% (15,36,57). The use of geometric mean or median IELT values is more representative of the treatment response in a population with this type of distribution (57,58). Furthermore, as a typical trial population has a broad range of baseline IELT values (0–60 s), reporting mean raw trial-end IELT may be misleading by incorrectly suggesting all subjects respond to that extent. The trial-end fold increase in geometric mean IELT compared to baseline is more representative of true treatment outcome and must be regarded as the contemporary universal standard for reporting IELT.

The reliability of stopwatch IELT alone in assigning PE status, the use of PROs to replace stopwatch IELT or the predictive value of single-item PRO measures compared to multiple-item measures are incompletely understood issues. PROs measures whilst providing important information, are at best subjective, relate to highly interpretable and imprecise constructs of ejaculation and their significance is weighted differently for different subjects. On the other hand, IELT may not adequately categorize subjects as some subjects with a brief IELT report little or no bother, are therefore asymptomatic and not “suffering” from PE. Clearly, none of the constructs of PE can universally distinguish subjects with PE from non-PE subjects. The current consensus is that a combination of stopwatch IELT and a validated, subject administered PRO inventory of control, satisfaction, personal distress and interpersonal distress can adequately identify PE status in prevalence

studies, in the screening phase of drug trials and measure response to treatment.

Assessment of baseline erectile and overall sexual function

The presence of comorbid ED and/or hypoactive sexual desire should be evaluated using a validated instrument such as the International Index of Erectile Function (IIEF) or the IIEF-5 (SHIM). Normal erectile function should be defined as an IIEF EF Domain ≥ 26 or IIEF-5 >21 (59,60). Recent data demonstrates that as many as half of subjects with ED also experience PE (35). In the European Premature Ejaculation Study (PEPA), ED present in 31.9% of men with PE compared to 11.8% of non-PE men (27). In the Global Study of Sexual Attitudes and Behaviours (GSSAB), the odds ratio for ED in men with PE ranged from was 6.0 in Europe and as high as 11.9 in South America (61). Consistent with this, ED is more prevalent in men with A-PE than L-PE (62). PE is also more common with increasing severity of ED after adjustment for age (63–65). Subjects with ED may either require higher levels of stimulation to achieve an erection or intentionally “rush” intercourse to prevent early detumescence of a partial erection, resulting in ejaculation with a brief latency (35). This may be compounded by the presence of high levels of performance anxiety related to their ED which serves to only worsen their prematurity. However, caution should be exercised in the diagnosis of comorbid ED in men with PE as 33.3% of potent men with PE confuse the ability to maintain erections prior to ejaculation and following ejaculation, record contradictory response/s to some/all questions of the SHIM especially Q3 and Q4 and receive a false positive SHIM diagnosis of ED (66). Although this is likely to limit subject recruitment in clinical trials by exclusion of subjects with low-range IELTs, it is unlikely to result in significantly different baseline IELTs or IELT distributions.

Treatment phase

In daily dosing SSRI studies, clinically significant ejaculatory delay may occur within 5–7 days but 5-HT_{1A} receptor desensitisation and maximal ejaculatory delay may not occur for 3–4 weeks (47). The treatment phase of daily dosing SSRI studies should have a duration of at least 6–8 weeks Although there is no clear data to identify the duration of treatment required to result in PRO response change, the anecdotal impression of the author derived from the clinical treatment of patients suggests that PRO change is likely to be a response to initial ejaculatory delay and may be delayed. As such, if PROs of control, satisfaction

and personal/interpersonal distress are included as trial outcomes, trial duration of at least 12 weeks is required to demonstrate statistically and clinically significant PRO response change. Trial visits should be scheduled at 4, 8 and 12 weeks (trial-end). In on-demand dosing studies using rapid acting, short life SSRIs, topical anaesthetics or other classes of drugs, a trial duration of at least 12 weeks is required to demonstrate significant PRO response change.

Administration of trial drug

Trial drug should be administered in accordance with the manufacturer's prescribing information. Although fixed dosing regimens provide important information about the relationship between dose, response and the incidence of adverse effects, a flexible dosing regimen which includes the opportunity to titrate the dose of trial or comparator drug to optimise efficacy and minimise treatment emergent adverse effects, provides additional efficacy data. The extent of subject compliance with dosing frequency can be assessed by audit of the subject diary and returned trial drug/s. On-demand SSRI studies should include days on which the drug is not taken, dependent of the half-life of the drug, to avoid inadvertent and unwanted 5-HT_{1A} receptor desensitisation and a daily dosing type response. Measurement of the time interval between dosing and intercourse, the drug-coital interval time (DCIT) is required to assess the relationship between time of dosing and efficacy (67).

Data collection and recording

Trial outcome data should be collected for each sexual event and recorded within 12 h in either a written diary or a trial-dedicated electronic PDA, PDA-phone or Smart-Phone diary. Electronic diary data can be either downloaded at the next trial visit or wireless transmitted to a central server computer within 12 h of each sexual event for archiving. ELT can be recorded with either a stopwatch operated by the partner or a blinded-stopwatch integrated into an electronic diary.

The diary should include the date/time of the onset of sexual activity, whether penetration was attempted, the date/time of that attempted penetration and whether penetration was achieved, global details of ejaculation, i.e., ante-portal, intra-vaginal/-anal/-oral, anejaculation, the intra-vaginal/-anal/-oral or masturbatory ejaculation latency time (IELT, AELT, OELT, MELT) in seconds and the responses to patient ± partner key single-item PROs (68).

Subject responses to self-administered multi-item multi-domain PE inventories and tertiary trial outcome measures are best administered at trial visits and responses recorded in a paper or electronic Clinical Research File (CRF).

Tertiary efficacy outcome measures are trial-specific and may include clinical global impression of change (CGIC), global efficacy, symptom severity impression and quality of life (QoL) assessments.

Post-trial phase

A 4-week treatment free post-trial phase is required to detect any residual adverse effects. Data is collected at either a scheduled final visit (16 weeks) or by investigator/research staff-initiated telephone contact.

Outcome measures

Patient reported outcome (PRO) measures

PROs are single-item diary questions or multi-item multi-domain questionnaires used as diagnostic tools and as intervention outcome measures to assess clinical improvement, intercourse-related subject and partner sexual satisfaction, relationship satisfaction, personal and interpersonal distress and subject and partner QoL. They can be assessed using validated single item questions, validated multi-item multi-domain PE inventories or validated omnibus sexual inventories (18,48,69-75). PROs for use in clinical trials of investigational drugs should conform to the guidelines of the relevant regulatory agency (76).

Standardized assessment measures such as validated questionnaires and PRO measures were developed primarily for use as research tools but are commonly used as an adjunct to a full medical/sexual history and self-estimation of ejaculatory latency in the evaluation of men presenting with self-reported PE. Some have shown good psychometric properties and are potentially valuable adjuncts for clinical screening and assessment.

Several PE measures have been described in the literature (48,50,51,72,73,77,78), although only a small number have undergone extensive psychometric testing and validation. Five validated questionnaires have been developed and published to date. Currently, there are two questionnaires that have extensive databases and meet most of the criteria for test development and validation: the premature ejaculation profile (PEP) and the index of premature ejaculation (IPE) (77,78). Two other measures, the Arabic and Chinese PE Questionnaires, have minimal validation or clinical trial data available and are not recommended for clinical use.

Premature ejaculation profile (PEP)

A four-item, self-report measure of premature ejaculation has been developed by Patrick *et al.* (78). The PEP is

comprised of single-item constructs of: (I) perceived control over ejaculation; (II) satisfaction with sexual intercourse; (III) personal distress related to ejaculation; (IV) interpersonal difficulty related to ejaculation and (V) an index or total score. Each of the four individual items is assessed on a 5-point scale, which are averaged to provide an index PE score. The measure has been used in observational studies and clinical trials of premature ejaculation (18). It has also been recommended for clinical use in evaluating the subjective components of the disorder. Validation studies have been performed in comparison to stopwatch measures of intra-vaginal latency and other PRO measures of sexual function and distress (18,78). The scale has adequate test-retest reliability (total scale =0.80) and moderate to strong correlations with stopwatch measured IELT. A major limitation of the scale is the lack of validated cut-off scores, which make it less suitable for use as a diagnostic or clinical screening tool. On the positive side, it is very brief and easy to administer and may be valuable for use in a clinical setting as a measure of treatment responsiveness.

Index of premature ejaculation (IPE)

The IPE was developed by Althof *et al.* (77). It is a 10 item self-administered questionnaire designed to evaluate sexual satisfaction, control and distress in men with premature ejaculation. It was developed using four stages: item pool development, initial psychometric analyses, patient interviews, and final psychometric analyses. The IPE contains three factor analytically derived domains: control, sexual satisfaction and distress. All three domains have shown adequate internal consistency and reliability, as well as known groups validity in comparing men with and without PE. Convergent validity against IELT was also strong for all three domains [control ($r=0.75$); sexual satisfaction ($r=0.60$) and distress ($r=0.68$)].

The IPE has the advantages of also being relatively brief and easy-to-administer, although the measure is not as brief as the PEP above. It also assesses clinically relevant domains and has adequate known groups' validity data. However, similar to the PEP, it lacks norms and diagnostic cut-offs, and has limited value as a diagnostic or screening measure for PE.

Depending on the specific need, the PEP or IPE are currently the preferred questionnaire measures for assessing PE, particularly when monitoring responsiveness to treatment. Overall, these measures may serve as useful adjuncts, but should not substitute for a detailed sexual history performed by a qualified clinician.

Omnibus sexual inventories

Omnibus sexual inventories such as the Golombok Russ

Inventory of Sexual Satisfaction (GRISS) (69) and Derogatis Sexual Function Inventory (DSFI) (79) are multi-item multi-domain questionnaires designed to assess global sexual function. The GRISS is a 12 domain, 28-item questionnaire which diagnoses the presence and severity of sexual dysfunction including PE in subjects and women and has no role in assessing treatment outcomes (69). The measure has good reliability and satisfactory validity but is more helpful with diagnosis than outcome. The DSFI comprises a 10 domain, 254-items questionnaire which is limited to assessing the extent of bother, frustration or distress in subjects with PE and has few questions on other aspects of PE such as latency or control (79). Both have demonstrated good reliability validity and sensitivity. Neither have good utility in PE intervention trials.

Global impression of change question

CGIC measures have high utility in clinical practice. The CGIC question (Has the treatment you have been taking improved your premature ejaculation?) asks patients to rate improvement or worsening of their PE compared with the start of the study using a 7-point response scale. It allows patients to overall evaluate their treatment response by self-interpretation of changes in ejaculatory latency, control, negative psychological consequences, sexual satisfaction and partner response. Higher CGIC ratings are correlated with greater improvement in latency, control, and satisfaction, and with greater reduction in distress, and interpersonal difficulty. The CGIC can provide clinicians in practice with a valid and brief outcome assessment of their patient's condition (80).

Responder definition

There is currently no published literature which identifies a meaningful and clinically significant threshold response to intervention. Statistical superiority to baseline or placebo outcome measures does not always imply a clinically significant response. Threshold response to intervention can be identified by either a threshold IELT fold increase or a composite PRO responder definition.

The point at which the IELT fold-increase achieved by intervention is associated with a significant reduction in personal distress probably represents a measure of intervention success. This data is currently not available but the author's anecdotal impression, derived from treatment of patients, suggests that a 3–4-fold-increase in IELT represents the threshold of intervention success with higher fold increases reflecting improving intervention success. Based upon a post-hoc path analysis of IELT and PRO data from the US observational trial, a composite PRO

responder definition of ≥ 2 -category increase in control and a 1-category decrease in personal distress from baseline (0–5 categorical scale) was developed and used as a efficacy outcome measure in a phase III Asia Pacific dapoxetine trial (18,81–83).

Adverse effects

By convention, adverse events are reported retrospectively by the subject at the next trial visit, and are recorded and rated by the investigator using MedDRA (Version 10.1) coding as “not related”, “doubtful”, “possible”, “probable” or “very likely related” to the trial drug (84). However, the next trial visit may not take place for up to 4 weeks, prompting concerns regarding the reliability of subject recall of the details of the adverse event, frequency, severity, duration and/or temporal relationship to trial drug/s dosing. Prospective reporting of adverse events within 24 h in a subject diary using a validated questionnaire, such as the UKU side effect rating scale for psychotropic/neuroleptic drugs, has been suggested (85,86).

Discontinuation rate

Withdrawal from a trial is usually due to lack of efficacy in both the placebo and active drug groups but may also relate to intolerable adverse effects. The rate of early withdrawals from studies varies with the type of trial. The inclusion of a placebo control raises the level of subject and investigator concern and increases withdrawal of subjects having an equivocal response from the trial.

The discontinuation rate due to lack of efficacy provides useful additional efficacy data which may be captured in a last observation carried forward (LOCF) analysis. It is particularly important to use the data from dropouts due to lack of efficacy if a trial has a high drop-out rate. Analysis of dropouts due to lack of efficacy should include separate analysis according to baseline PE severity of the subjects. A further measure of interventional efficacy is provided by a survival analysis comparing the time to withdrawal due to lack of efficacy.

Partner outcome measures (PaROs)

The inability to control and defer ejaculation until the female partner was sexually satisfied on at least 50% of intercourse attempts was proposed as a definition of PE by Masters and Johnson (6). Although an inherent problem exists in defining a man as dysfunctional based on the sexual responsiveness of his partner, several studies have reported that the effects of PE on the partner are integral to understanding the impact of PE on the male and on the sexual relationship as a whole (87–89). The extent of

psychological impact on patients, partners and the overall relationship are perhaps the most important aspect of treatment seeking behavior and best define the severity of PE. The report that subjects with PE regard fulfilling a partner's needs as the most important factor contributing to an overall sense of sexual satisfaction is testament to the pivotal role of the partner in defining outcome success (90).

There is, however, limited information regarding the effect of PE on the partner. Research in partners of men with erectile dysfunction suggests that a woman's sexual difficulties can be contingent on her partner's sexual dysfunction. Hobbs *et al* reported that 77.7% of PE partners had at least one sexual dysfunction, compared to 42.7% of the control group (91). Patrick *et al.* reported that 44% of partners of subjects with PE rated their extent of personal distress as “quite a bit” or “extremely” compared to 3% in a group of partners of normal controls (18). Patrick *et al.* also reported that partner PRO measures differentiated subjects with PE from subjects without PE and correlated moderately with measures of IELT and subject PRO measures. Furthermore, Rosen *et al.* report that partner distress was one of several factors which were more influential in determining PE status than IELT (56). However, partner perceptions of PE generally indicated less dysfunction than those of subjects (18). Although PE adversely affects partner sexual satisfaction, it appears to have minimal impact upon relationship satisfaction (88). Furthermore, partners of subjects with PE report relatively high levels of female sexual dysfunction (92,93). The observation that PE often pre-dates the time of onset of the women's sexual symptoms, suggests that PE may be a risk factor for female sexual dysfunction (92).

If PE is to be regarded as a disorder that affects both subjects and their partners, PaROs must be regarded as important measures in determining PE severity and treatment outcomes. The DSM-IV-TR, World Health Organization (WHO) ICD-10, World Health Organization (WHO) 2nd International Consultation on Sexual Health and American Urological Society (AUA) multivariate definitions of PE include partner distress and/or sexual satisfaction as constructs of PE (7,28,29,31). PE questionnaires used in observational and as secondary outcome measures in intervention studies should include specific sub-scales for PaROs. Although 3 of the 4 currently available PE treatment outcome questionnaires include questions on PaROs (71,72,75), only Althof and Corty's PE index (PEI) has a specific partner questionnaire (75). Although the IPE is regarded as the state-of-the-art

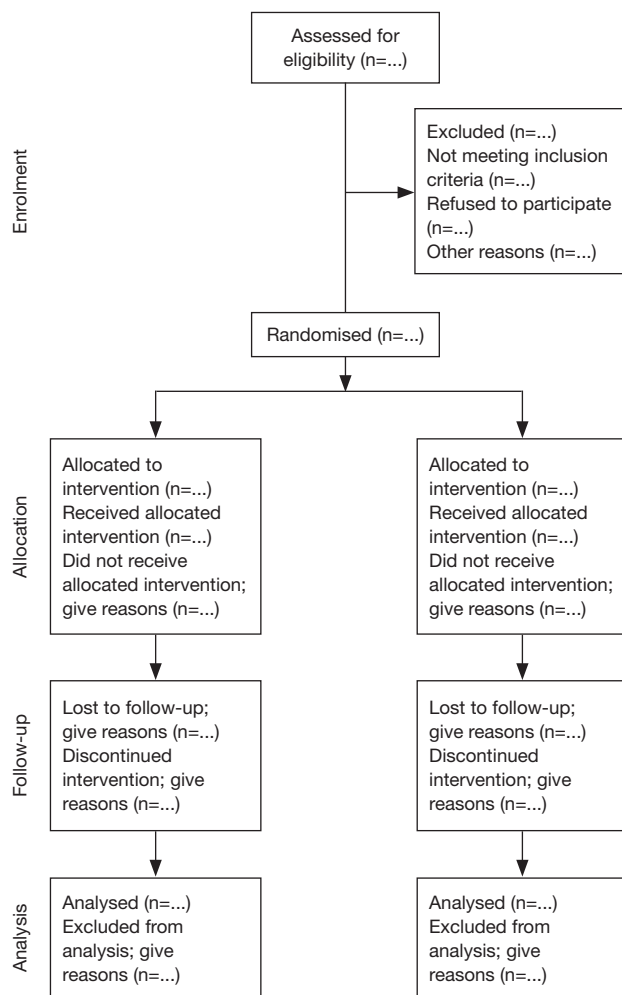


Figure 2 Flow diagram depicting progress of subjects through the four stages of a RCT (105).

PE questionnaire, it is limited by the lack of a partner subscale (94). Until such time that a validated PE questionnaire with a partner subscale is available, validated omnibus sexual inventories such as the Golombok Rust Inventory of Sexual Satisfaction (GRISS) or the Derogatis Interview for Sexual Functioning (DISF/DISF-SR) (69,95), specific female sexual dysfunction (FSD) scales such as the Female Distress Scale (FSDS), the Female Sexual Function Index (FSFI), the Brief Index of Sexual Functioning for Women (BISF-W), the Sexual Function Questionnaire (SFQ), the Changes in Sexual Functioning Questionnaire (CSFQ), the McCoy Female Sexuality Questionnaire (MFSQ), the Sexual Satisfaction and Distress Scale (SSS-W) (96-102), or specific distress scales such as Hospital Anxiety and Depression Scale (HADS) (103) can be used as

surrogate PE partner distress and satisfaction scales.

Reporting clinical trials

If a reader is to understand the conduct of a RCT and to assess the validity of its results, authors must convey with total transparency full details of trial design, conduct, analysis, and interpretation. The Consolidated Standards of Reporting Trials (CONSORT) statement was developed by a group of clinical investigators, statisticians, epidemiologists, and biomedical journal editors to help authors improve reporting by the use of a checklist and flow diagram (104). New evidence and responses to several criticisms of the original statement have been incorporated in a revised CONSORT statement (105). CONSORT has been supported by a growing number of medical and health-care journals and editorial groups, including the International Committee of Medical Journal Editors (ICMJE), The Vancouver Group, the Council of Science Editors (CSE), and the World Association of Medical Editors (WAME).

The checklist items relate to the content of the Title, Abstract, Introduction, Methods, Results, and Discussion. These inclusion of these items was based upon empirical evidence suggesting that failure to report this information is associated with biased estimates of treatment effect, or because the information is essential to judge the reliability or relevance of the findings. The flow chart depicts the passage of subjects through the four stages of a RCT (enrolment, intervention allocation, follow-up, and analysis) and shows the number of participants, for each intervention group, included in the primary data analysis (Figure 2). The checklist and flow chart are primarily intended for use in writing, reviewing, or assessing reports of simple two-group parallel RCTs. Inadequate reporting makes the interpretation of RCTs difficult, if not impossible and results cannot be generalised to the general trial population. Potentially, the use of CONSORT should positively influence the manner in which RCTs are conducted and reduce or eliminate inadequate reporting of RCTs (106,107).

Conclusions

Data from PE intervention studies are only reliable, interpretable and capable of being generalised to patients with PE when derived from prospective, double-blind, placebo-controlled RCTs. The study population should be

defined by populations defined by the ISSM definition of PE. The ISSM definition of PE reflects the contemporary understanding of PE and represents the state-of-the-art definition of PE and is recommended as the basis of diagnosis of PE for all PE clinical trials (38). Study endpoints should include ELT and PROs/PaROs measures of perceived ejaculatory control and personal/partner/relationship distress determined by single-item questions or multi-item questionnaires. There is no current data which identifies a clinically significant threshold response to intervention and statistical superiority to baseline or placebo outcome measures does not always imply a clinically significant response.

Acknowledgements

None.

Footnote

Conflicts of Interest: Dr. McMahon is or has been a consultant, investigator and speaker for Johnson & Johnson, Janssen Cilag, Menarini, Ixchelsis, Absorption Pharmaceuticals, NeuroHealing and Plethora.

References

- Olivier B, van Oorschot R, Waldinger MD. Serotonin, serotonergic receptors, selective serotonin reuptake inhibitors and sexual behaviour. *Int Clin Psychopharmacol* 1998;13 Suppl 6:S9-14.
- Waldinger MD, Zwinderman AH, Schweitzer DH, et al. Relevance of methodological design for the interpretation of efficacy of drug treatment of premature ejaculation: a systematic review and meta-analysis. *Int J Impot Res* 2004;16:369-81.
- Waldinger MD, Hengeveld MW, Zwinderman AH. Paroxetine treatment of premature ejaculation: a double-blind, randomized, placebo-controlled study. *Am J Psychiatry* 1994;151:1377-9.
- McMahon CG, Touma K. Treatment of premature ejaculation with paroxetine hydrochloride as needed: 2 single-blind placebo controlled crossover studies. *J Urol* 1999;161:1826-30.
- Semans JH. Premature ejaculation: a new approach. *South Med J* 1956;49:353-8.
- Masters WH, Johnson VE. *Human Sexual Inadequacy*. Boston: Little Brown; 1970:92-115.
- McMahon CG, Meston C, Waldinger MD, et al. Disorders of orgasm in men and women, ejaculatory disorders in men. In: Lue TF, Basson R, Rosen R, et al. editors. *Sexual medicine: sexual dysfunctions in men and women*. Paris: Health Publications, 2004.
- International Committee of Harmonization (ICH). *ICH Harmonised Tripartite Guideline for Good Clinical Practice*. Surrey: Brookwood Medical Publications, 1996.
- Hutchinson DR. *The trial investigator's GCP handbook: a practical guide to ICH requirements (Clinical trials)*. Surrey: Brookwood Medical Publications, 1997.
- World Medical Association. *Helsinki declaration of human clinical trials*. Available online: <http://www.wma.net/en/30publications/10policies/b3/>. Accessed 2016.
- Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998;59 Suppl 20:22-33;quiz 34-57.
- Shapiro B. Premature ejaculation: a review of 1130 cases. *J Urol* 1943;50:6.
- Godpodinoff ML. Premature ejaculation: clinical subgroups and etiology. *J Sex Marital Ther* 1989;15:130-4.
- Waldinger MD, Hengeveld MW, Zwinderman AH, et al. An empirical operationalization study of DSM-IV diagnostic criteria for premature ejaculation. *Int J Psychiatry Clin Pract* 1998;2:287-93.
- McMahon CG. Long term results of treatment of premature ejaculation with selective serotonin re-uptake inhibitors. *Int J Impot Res* 2002;14:S19.
- Waldinger MD, McIntosh J, Schweitzer DH. A five-nation survey to assess the distribution of the intravaginal ejaculatory latency time among the general male population. *J Sex Med* 2009;6:2888-95.
- Waldinger MD, Quinn P, Dilleen M, et al. A multinational population survey of intravaginal ejaculation latency time. *J Sex Med* 2005;2:492-7.
- Patrick DL, Althof SE, Pryor JL, et al. Premature ejaculation: an observational study of men and their partners. *J Sex Med* 2005;2:358-67.
- Waldinger MD, Schweitzer DH. The use of old and recent DSM definitions of premature ejaculation in observational studies: a contribution to the present debate for a new classification of PE in the DSM-V. *J Sex Med* 2008;5:1079-87.
- Waldinger MD, Schweitzer DH. Changing paradigms from a historical DSM-III and DSM-IV view toward an evidence-based definition of premature ejaculation.

- Part II--proposals for DSM-V and ICD-11. *J Sex Med* 2006;3:693-705.
21. Waldinger MD. History of premature ejaculation. In: Jannini EA, McMahon CG, Waldinger MD, editors. *Premature ejaculation: from etiology to diagnosis and treatment*. New York: Springer, 2013:5-24.
 22. Serefoglu EC, Cimen HI, Atmaca AF, et al. The distribution of patients who seek treatment for the complaint of ejaculating prematurely according to the four premature ejaculation syndromes. *J Sex Med* 2010;7:810-5.
 23. Serefoglu EC, Yaman O, Cayan S, et al. Prevalence of the complaint of ejaculating prematurely and the four premature ejaculation syndromes: results from the Turkish Society of Andrology Sexual Health Survey. *J Sex Med* 2011;8:540-8.
 24. Zhang X, Gao J, Liu J, et al. Distribution and factors associated with four premature ejaculation syndromes in outpatients complaining of ejaculating prematurely. *J Sex Med* 2013;10:1603-11.
 25. Gao J, Zhang X, Su P, et al. Prevalence and factors associated with the complaint of premature ejaculation and the four premature ejaculation syndromes: a large observational study in China. *J Sex Med* 2013;10:1874-81.
 26. Waldinger MD. Premature ejaculation: definition and drug treatment. *Drugs* 2007;67:547-68.
 27. Porst H, Montorsi F, Rosen RC, et al. The Premature Ejaculation Prevalence and Attitudes (PEPA) survey: prevalence, comorbidities, and professional help-seeking. *Eur Urol* 2007;51:816-23; discussion 824.
 28. *Diagnostic And Statistical Manual of Mental Disorders, DSM-IV 4th*. Washington D.C.: American Psychiatric Association, 1994:509-11.
 29. World Health Organization. *International Classification of Diseases and Related Health Problems (10th ed.)*. Geneva: World Health Organization, 1994.
 30. Metz M, McCarthy B. *Coping with premature ejaculation: How to overcome PE, please your partner and have great sex*. Oakland, CA: New Harbinger Publications, 2004.
 31. Montague DK, Jarow J, Broderick GA, et al. AUA guideline on the pharmacologic management of premature ejaculation. *J Urol* 2004;172:290-4.
 32. Colpi G, Weidner W, Jungwirth A, et al. EAU guidelines on ejaculatory dysfunction. *Eur Urol* 2004;46:555-8.
 33. McMahon CG, Abdo C, Incrocci I, et al. Disorders of orgasm and ejaculation in men. In: Lue TF, Basson R, Rosen R, et al. editors. *Sexual medicine: sexual dysfunctions in men and women (2nd International Consultation on Sexual Dysfunctions-Paris)*. Paris, France: Health Publications, 2004:409-68.
 34. Waldinger MD, Zwinderman AH, Olivier B, et al. Proposal for a definition of lifelong premature ejaculation based on epidemiological stopwatch data. *J Sex Med* 2005;2:498-507.
 35. Jannini EA, Lombardo F, Lenzi A. Correlation between ejaculatory and erectile dysfunction. *Int J Androl* 2005;28 Suppl 2:40-5.
 36. Giuliano F, Patrick DL, Porst H, et al. Premature ejaculation: results from a five-country European observational study. *Eur Urol* 2008;53:1048-57.
 37. McMahon CG, Althof S, Waldinger MD, et al. An evidence-based definition of lifelong premature ejaculation: report of the International Society for Sexual Medicine Ad Hoc Committee for the Definition of Premature Ejaculation. *BJU Int* 2008;102:338-50.
 38. Serefoglu EC, McMahon CG, Waldinger MD, et al. An evidence-based unified definition of lifelong and acquired premature ejaculation: report of the second international society for sexual medicine ad hoc committee for the definition of premature ejaculation. *Sex Med* 2014;2:41-59.
 39. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (5th edition) (DSM-5)*. Washington, DC: American Psychiatric Association, 2013.
 40. Bodenheimer T. Uneasy alliance--clinical investigators and the pharmaceutical industry. *N Engl J Med* 2000;342:1539-44.
 41. Tomaszewska W, Peselow ED, Barouche F, et al. Antecedent life events, social supports and response to antidepressants in depressed patients. *Acta Psychiatr Scand* 1996;94:352-7.
 42. Krupnick JL, Sotsky SM, Simmens S, et al. The role of the therapeutic alliance in psychotherapy and pharmacotherapy outcome: findings in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *J Consult Clin Psychol* 1996;64:532-9.
 43. Kraemer HC, Wilson GT, Fairburn CG, et al. Mediators and moderators of treatment effects in randomized clinical trials. *Arch Gen Psychiatry* 2002;59:877-83.
 44. Baxter LR Jr, Schwartz JM, Bergman KS, et al. Caudate glucose metabolic rate changes with both drug and behavior therapy for obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:681-9.
 45. Kandel ER. A new intellectual framework for psychiatry. *Am J Psychiatry* 1998;155:457-69.
 46. Kandel ER. Biology and the future of psychoanalysis: a new intellectual framework for psychiatry revisited. *Am J*

- Psychiatry 1999;156:505-24.
47. Waldinger MD, Berendsen HH, Blok BF, et al. Premature ejaculation and serotonergic antidepressants-induced delayed ejaculation: the involvement of the serotonergic system. *Behav Brain Res* 1998;92:111-8.
 48. Symonds T, Perelman MA, Althof S, et al. Development and validation of a premature ejaculation diagnostic tool. *Eur Urol* 2007;52:565-73.
 49. Fleiss J. *The design and analysis of clinical experiments*. New York, NY: John Wiley & Sons Inc; 1999:49-51.
 50. Symonds T, Perelman M, Althof S, et al. Further evidence of the reliability and validity of the premature ejaculation diagnostic tool. *Int J Impot Res* 2007;19:521-5.
 51. Serefoglu EC, Cimen HI, Ozdemir AT, et al. Turkish validation of the premature ejaculation diagnostic tool and its association with intravaginal ejaculatory latency time. *Int J Impot Res* 2009;21:139-44.
 52. Serefoglu EC, Yaman O, Cayan S, et al. The comparison of premature ejaculation assessment questionnaires and their sensitivity for the four premature ejaculation syndromes: results from the Turkish society of andrology sexual health survey. *J Sex Med* 2011;8:1177-85.
 53. Kam SC, Han DH, Lee SW. The diagnostic value of the premature ejaculation diagnostic tool and its association with intravaginal ejaculatory latency time. *J Sex Med* 2011;8:865-71.
 54. Althof SE, Levine SB, Corty EW, et al. A double-blind crossover trial of clomipramine for rapid ejaculation in 15 couples. *J Clin Psychiatry* 1995;56:402-7.
 55. Pryor JL, Broderick GA, Ho KF, et al. Comparison of estimated versus measured intravaginal ejaculatory latency time (IELT) in men with and without premature ejaculation (PE). *J Sex Med* 2005;3:54(abstract 126).
 56. Rosen RC, McMahon CG, Niederberger C, et al. Correlates to the clinical diagnosis of premature ejaculation: results from a large observational study of men and their partners. *J Urol* 2007;177:1059-64; discussion 1064.
 57. Waldinger MD, Hengeveld MW, Zwinderman AH, et al. Effect of SSRI antidepressants on ejaculation: a double-blind, randomized, placebo-controlled study with fluoxetine, fluvoxamine, paroxetine, and sertraline. *J Clin Psychopharmacol* 1998;18:274-81.
 58. Waldinger MD, Zwinderman AH, Olivier B, et al. Geometric mean IELT and premature ejaculation: appropriate statistics to avoid overestimation of treatment efficacy. *J Sex Med* 2008;5:492-9.
 59. Rosen RC, Riley A, Wagner G, et al. The international index of erectile function (IIEF): a multidimensional scale for assessment of erectile dysfunction. *Urology* 1997;49:822-30.
 60. Cappelleri JC, Siegel RL, Glasser DB, et al. Relationship between patient self-assessment of erectile dysfunction and the sexual health inventory for men. *Clin Ther* 2001;23:1707-19.
 61. Laumann EO, Nicolosi A, Glasser DB, et al. Sexual problems among women and men aged 40-80 y: prevalence and correlates identified in the Global Study of Sexual Attitudes and Behaviors. *Int J Impot Res* 2005;17:39-57.
 62. Basile Fasolo C, Mirone V, Gentile V, et al. Premature ejaculation: prevalence and associated conditions in a sample of 12,558 men attending the andrology prevention week 2001--a study of the Italian Society of Andrology (SIA). *J Sex Med* 2005;2:376-82.
 63. Corona G, Petrone L, Mannucci E, et al. Psychobiological correlates of rapid ejaculation in patients attending an andrologic unit for sexual dysfunctions. *Eur Urol* 2004;46:615-22.
 64. El-Sakka AI. Association of risk factors and medical comorbidities with male sexual dysfunctions. *J Sex Med* 2007;4:1691-700.
 65. el-Sakka AI. Severity of erectile dysfunction at presentation: effect of premature ejaculation and low desire. *Urology* 2008;71:94-8.
 66. McMahon CG. Screening for erectile dysfunction in men with lifelong premature ejaculation--Is the Sexual Health Inventory for Men (SHIM) reliable? *J Sex Med* 2009;6:567-73.
 67. Waldinger MD, Zwinderman AH, Olivier B. On-demand treatment of premature ejaculation with clomipramine and paroxetine: a randomized, double-blind fixed-dose study with stopwatch assessment. *Eur Urol* 2004;46:510-5; discussion 516.
 68. Waldinger MD. Four measures of investigating ejaculatory performance. *J Sex Med* 2007;4:520.
 69. Rust J, Golombok S. The GRISS: a psychometric instrument for the assessment of sexual dysfunction. *Arch Sex Behav* 1986;15:157-65.
 70. Lee HS, Song DH, Kim CH, et al. An open clinical trial of fluoxetine in the treatment of premature ejaculation. *J Clin Psychopharmacol* 1996;16:379-82.
 71. Hartmann U. The "PEQUEST": a multidimensional instrument for the assessment of premature ejaculation. *Int J Impot Res* 1996;8:119.
 72. Yuan YM, Xin ZC, Jiang H, et al. Sexual function of premature ejaculation patients assayed with Chinese Index

- of Premature Ejaculation. *Asian J Androl* 2004;6:121-6.
73. Arafa M, Shamloul R. Development and evaluation of the Arabic Index of Premature Ejaculation (AIPE). *J Sex Med* 2007;4:1750-6.
 74. Abdo CH. The male sexual quotient: a brief, self-administered questionnaire to assess male sexual satisfaction. *J Sex Med* 2007;4:382-9.
 75. Althof SE, Corty EW. Pentech Ejaculation Inventory (personal communication, 2000).
 76. US Food and Drug Administration (FDA). Guidance for industry, patient reported outcome measures: use in medical product development to support labeling claims. February 2006. Available online: <http://www.fda.gov/downloads/Drugs/.../Guidances/UCM193282.pdf>
 77. Althof S, Rosen R, Symonds T, et al. Development and validation of a new questionnaire to assess sexual satisfaction, control, and distress associated with premature ejaculation. *J Sex Med* 2006;3:465-75.
 78. Patrick DL, Giuliano F, Ho KF, et al. The Premature Ejaculation Profile: validation of self-reported outcome measures for research and practice. *BJU Int* 2009;103:358-64.
 79. Derogatis LR, Melisaratos N. The DSFI: a multidimensional measure of sexual functioning. *J Sex Marital Ther* 1979;5:244-81.
 80. Althof SE, Brock GB, Rosen RC, et al. Validity of the patient-reported Clinical Global Impression of Change as a measure of treatment response in men with premature ejaculation. *J Sex Med* 2010;7:2243-52.
 81. Patrick DL, Rowland D, Rothman M. Interrelationships among measures of premature ejaculation: the central role of perceived control. *J Sex Med* 2007;4:780-8.
 82. McMahon CG, Park NC, Zhao Y, et al. Treatment of premature ejaculation in the Asia-Pacific Region: results from a phase III double-blind, parallel-group study of dapoxetine. Jeju, Korea: Asia Pacific Society for Sexual Medicine (APSSM); October 6-10, 2007.
 83. Shabsigh R, Patrick DL, Rowland DL, et al. Perceived control over ejaculation is central to treatment benefit in men with premature ejaculation: results from phase III trials with dapoxetine. *BJU Int* 2008;102:824-8.
 84. MedDRA. MedDRA - medical dictionary for regulatory activities (Ver 19.1). Available online: <http://www.meddra.org/>. Accessed 2016.
 85. Lingjaerde O, Ahlfors UG, Bech P, et al. The UKU side effect rating scale. A new comprehensive rating scale for psychotropic drugs and a cross-sectional study of side effects in neuroleptic-treated patients. *Acta Psychiatr Scand Suppl* 1987;334:1-100.
 86. Waldinger MD, Schweitzer DH, Olivier B. Dapoxetine treatment of premature ejaculation. *Lancet* 2006;368:1869; author reply 1869-70.
 87. Symonds T, Roblin D, Hart K, et al. How does premature ejaculation impact a man's life? *J Sex Marital Ther* 2003;29:361-70.
 88. Byers ES, Grenier G. Premature or rapid ejaculation: heterosexual couples' perceptions of men's ejaculatory behavior. *Arch Sex Behav* 2003;32:261-70.
 89. Metz ME, Pryor JL. Premature ejaculation: a psychophysiological approach for assessment and management. *J Sex Marital Ther* 2000;26:293-320.
 90. Rowland D, Perelman M, Althof S, et al. Self-reported premature ejaculation and aspects of sexual functioning and satisfaction. *J Sex Med* 2004;1:225-32.
 91. Hobbs K, Symonds T, Abraham L, et al. Sexual dysfunction in partners of men with premature ejaculation. *Int J Impot Res* 2008;20:512-7.
 92. Riley A, Riley E. Premature ejaculation: presentation and associations. An audit of patients attending a sexual problems clinic. *Int J Clin Pract* 2005;59:1482-7.
 93. Yáñez D, Castelo-Branco C, Hidalgo LA, et al. Sexual dysfunction and related risk factors in a cohort of middle-aged Ecuadorian women. *J Obstet Gynaecol* 2006;26:682-6.
 94. Symonds T, Althof SE, Rosen RC, et al. Questionnaire assessment of ejaculatory control: development and validation of a new instrument. *Int J Imp Res* 2002;14:S33(abstract PS-7-1).
 95. Derogatis LR. The Derogatis Interview for Sexual Functioning (DISF/DISF-SR): an introductory report. *J Sex Marital Ther* 1997;23:291-304.
 96. Derogatis L, Rust J, Golombok S, et al. Validation of the profile of female sexual function (PFSF) in surgically and naturally menopausal women. *J Sex Marital Ther* 2004;30:25-36.
 97. Meston C, Trapnell P. Development and validation of a five-factor sexual satisfaction and distress scale for women: the Sexual Satisfaction Scale for Women (SSS-W). *J Sex Med* 2005;2:66-81.
 98. McCoy NL. The McCoy female sexuality questionnaire. *Qual Life Res* 2000;9:739-45.
 99. Clayton AH, McGarvey EL, Clavet GJ. The Changes in Sexual Functioning Questionnaire (CSFQ): development, reliability, and validity. *Psychopharmacol Bull* 1997;33:731-45.

100. Quirk FH, Heiman JR, Rosen RC, et al. Development of a sexual function questionnaire for clinical trials of female sexual dysfunction. *J Womens Health Gend Based Med* 2002;11:277-89.
101. Taylor JF, Rosen RC, Leiblum SR. Self-report assessment of female sexual function: psychometric evaluation of the Brief Index of Sexual Functioning for Women. *Arch Sex Behav* 1994;23:627-43.
102. Rosen R, Brown C, Heiman J, et al. The Female Sexual Function Index (FSFI): a multidimensional self-report instrument for the assessment of female sexual function. *J Sex Marital Ther* 2000;26:191-208.
103. Spinhoven P, Ormel J, Sloekers PP, et al. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psychol Med* 1997;27:363-70.
104. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996;276:637-9.
105. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001;357:1191-4.
106. Moher D, Jones A, Lepage L, et al. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA* 2001;285:1992-5.
107. Egger M, Jüni P, Bartlett C, et al. Value of flow diagrams in reports of randomized controlled trials. *JAMA* 2001;285:1996-9.

Cite this article as: McMahon CG. The design and methodology of premature ejaculation interventional studies. *Transl Androl Urol* 2016;5(4):508-525. doi: 10.21037/tau.2016.03.28