



STUDY PROTOCOL

MANTA and MANTA-RAY: Rationale and Design of Trials Evaluating Effects of Filgotinib on Semen Parameters in Patients with Inflammatory Diseases

Wayne J. G. Hellstrom · Radboud J. E. M. Dolhain · Timothy E. Ritter · Timothy R. Watkins · Sarah J. Arterburn · Goele Dekkers · Angi Gillen · Caroline Tonussi · Leen Gilles · Alessandra Oortwijn · Katrien Van Beneden · Dick E. de Vries · Suresh C. Sikka · Dirk Vanderschueren · Walter Reinisch

Received: January 21, 2022 / Accepted: April 14, 2022 / Published online: May 25, 2022
© The Author(s) 2022

ABSTRACT

Introduction: The phase 2 MANTA and MANTA-RAY studies were developed in consultation with global regulatory authorities to investigate potential impacts of filgotinib, a Janus kinase 1 preferential inhibitor, on semen parameters in men with active inflammatory diseases. Here we describe the methods and rationale for these studies.

Methods and Rationale: The MANTA and MANTA-RAY studies included men (aged 21–65 years) with active inflammatory bowel disease (IBD) and rheumatic diseases,

respectively. Participants had no history of reproductive health issues, and the following semen parameter values (\geq 5th percentile of World Health Organization reference values) at baseline: semen volume \geq 1.5 mL, total sperm/ejaculate \geq 39 million, sperm concentration \geq 15 million/mL, sperm total motility \geq 40% and normal sperm morphology \geq 30%. Each trial included a 13-week, randomized, double-blind, placebo-controlled period (filgotinib 200 mg vs placebo, up to $N = 125$ per arm), for pooled analysis of the week-13 primary endpoint (proportion of participants with \geq 50% decrease from baseline in sperm concentration). All semen assessments were based on two samples (\leq 14 days apart) to minimize effects of physiological variation; stringent standardization processes were applied across assessment

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12325-022-02168-4>.

W. J. G. Hellstrom (✉)
Tulane Health Sciences Center, New Orleans, LA, USA
e-mail: whellst@tulane.edu

R. J. E. M. Dolhain
Erasmus University Medical Center Rotterdam, Rotterdam, Netherlands

T. E. Ritter
GI Alliance, Southlake, TX, USA

T. R. Watkins · S. J. Arterburn · A. Gillen · C. Tonussi
Gilead Sciences, Inc., Foster City, CA, USA

G. Dekkers · L. Gilles · K. Van Beneden
Galapagos NV, Mechelen, Belgium

A. Oortwijn · D. E. de Vries
Galapagos BV, Leiden, Netherlands

S. C. Sikka
Tulane University School of Medicine, New Orleans, LA, USA

D. Vanderschueren
Laboratory of Experimental and Clinical Endocrinology, KU Leuven, Leuven, Belgium

W. Reinisch
Medical University of Vienna, Vienna, Austria

sites. From week 13, MANTA and MANTA-RAY study designs deviated owing to disease-specific considerations. All subjects with a $\geq 50\%$ decrease in sperm parameters continued the study in the monitoring phase until reversibility, or up to a maximum of 52 weeks, with standard of care as treatment. Overall conclusions from MANTA and MANTA-RAY will be based on the totality of the data, including secondary/exploratory measures (e.g. sperm motility/morphology, sex hormones, reversibility of any effects on semen parameters).

Conclusions: Despite the complexities, the MANTA and MANTA-RAY studies form a robust trial programme that is the first large-scale, placebo-controlled evaluation of potential impacts of an advanced IBD and rheumatic disease therapy on semen parameters.

Trial Registration: EudraCT numbers 2017-000402-38 and 2018-003933-14; ClinicalTrials.gov identifiers NCT03201445 and NCT03926195.

PLAIN LANGUAGE SUMMARY

Filgotinib is a treatment for patients with ulcerative colitis and rheumatoid arthritis, and is being studied in other inflammatory diseases. Filgotinib works by blocking Janus kinase 1, an intracellular protein involved in inflammatory signalling processes. We designed the MANTA and MANTA-RAY trials with global health agencies to find out if filgotinib decreases the

quality of semen in men with active inflammatory bowel disease (ulcerative colitis or Crohn's disease) (MANTA) or rheumatic disease (rheumatoid arthritis, psoriatic arthritis, ankylosing spondylitis or non-radiographic axial spondylitis) (MANTA-RAY). This paper describes the design of the two trials.

Patients had normal sperm measurements and could not have had previous reproductive health issues. Nearly 250 patients were included in each trial. In both MANTA and MANTA-RAY, half of the patients were treated with 200 mg of filgotinib once a day for 13 weeks, and the other half with placebo. We determined if any patients had a decrease in number of sperm cells per millilitre (sperm concentration) by at least half after 13 weeks of treatment. We then monitored any patients who had such a decrease in sperm concentration for up to 52 weeks (while they received standard of care treatment) or until the decrease was reversed.

The conclusions from the trials will be in a different paper and will be based on all the final data, including changes in sex hormones. This is the first large-scale clinical trial programme to measure the effect of a treatment on sperm in men with inflammatory bowel disease or rheumatic diseases.

Keywords: Filgotinib; MANTA; MANTA-RAY; Inflammatory bowel disease; Janus kinase 1 preferential inhibitor; Randomized controlled trials; Rheumatoid diseases; Semen parameters; Reproductive health

Key Summary Points

Filgotinib is a once-daily oral Janus kinase 1 preferential inhibitor, approved for the treatment of rheumatoid arthritis in Europe, the UK and Japan, and for ulcerative colitis treatment in Europe, and is under investigation for the treatment of Crohn's disease.

Preclinical studies of filgotinib in rats and dogs demonstrated histopathological lesions (germ cell depletion/degeneration and/or tubular vacuolation) in testes with correlating findings in epididymides; the highest exposure level with no adverse effects on the testis in dogs (the most sensitive species) was twofold above the exposure seen in humans taking filgotinib 200 mg once daily, the highest dose evaluated in phase 3 studies.

The randomized, double-blind, placebo-controlled, phase 2 MANTA and MANTA-RAy studies were designed, upon consultation with global regulatory authorities, to evaluate the impact (if any) of filgotinib 200 mg on semen parameters in participants with active inflammatory bowel disease (IBD) and rheumatic diseases, respectively, and are the first large studies to assess the effects of an advanced therapy for IBD and rheumatoid diseases on semen parameters.

Despite several challenges, such as stringent selection criteria (active inflammatory disease, no history of reproductive health issues, semen parameters \geq 5th percentile of World Health Organization reference values) with the potential to slow recruitment, and logistical complexities (e.g. standardization of assessments of semen parameters across multiple, international sites), a unique and robust trial programme was successfully executed.

The primary endpoint (combined across both trials) was the proportion of participants with a \geq 50% reduction from baseline in sperm concentration after 13 weeks of treatment with filgotinib 200 mg versus placebo; overall conclusions, however, will be based on the totality of the data, including secondary and exploratory endpoints.

INTRODUCTION

Chronic inflammatory diseases are increasing in prevalence [1, 2], probably owing to changes in social, environmental and lifestyle factors [3]. The classic signs, symptoms and complications of chronic inflammatory conditions such as inflammatory bowel disease (IBD) and rheumatoid arthritis (RA) are well known [4, 5]. Less widely appreciated are the negative impacts of chronic inflammatory conditions on sexual and reproductive health [6–16].

Despite advances in the management of chronic inflammatory diseases, especially with regard to the introduction of biologic agents, significant unmet treatment needs still remain in terms of rates and durability of response, tolerability and safety [17–19]. Janus kinase (JAK) inhibitors have been investigated and approved as treatments for inflammatory diseases owing to the wide range of cytokine pathways regulated via JAK/signal transducer and activator of transcription signalling [20]. Filgotinib is an orally administered, adenosine triphosphate-competitive, reversible inhibitor of the JAK family, with preferential inhibitory activity against JAK1 [21]. After oral administration, filgotinib is quickly metabolized by carboxylesterase 2 to GS-829845, which has similar properties of JAK1 preferential inhibition, but a longer half-life than filgotinib [22, 23]. Consequently, GS-829845 exposure is 16–20-fold higher than filgotinib exposure, and thus has the greatest contribution to the overall pharmacodynamic effects of filgotinib treatment [22].

Filgotinib (100 mg or 200 mg once daily) is approved in Europe and the UK as monotherapy, or in combination with methotrexate, for the treatment of moderately to severely active RA in patients with inadequate response or intolerance to one or more disease-modifying anti-rheumatic drugs (DMARDs) [24], and in Japan for the treatment of RA (including prevention of structural joint damage) in patients who have had inadequate response to conventional therapies. Filgotinib has also recently been approved in Europe for the treatment of moderately to severely active ulcerative colitis (UC), on the basis of positive findings from the multicentre, randomized, double-blind, placebo-controlled phase 2b/3 SELECTION study [25]. Filgotinib is also under investigation in patients with Crohn's disease (CD) in a global, phase 3 clinical trial programme (the DIVERSITY1 study; EudraCT number 2016-001367-36; ClinicalTrials.gov identifier NCT02914561).

Alongside clinical trials, the development of new therapeutic compounds requires extensive *in vitro* and *in vivo* testing. Evaluation of potential testicular effects is a standard component of this process and is routinely assessed during repeat-dose toxicity studies. Preclinical, chronic toxicity studies in rats and dogs demonstrated that male reproductive organs were affected by filgotinib, but not by its primary metabolite, GS-829845 [26]. These effects appeared at different exposure levels across species, with dogs being the most sensitive [26]. The observed lesions comprised germ cell depletion/degeneration and/or tubular vacuolation in testes with correlating findings in epididymides (reduced sperm count and/or cell debris) at high doses of filgotinib [26]. In rats, similar changes were observed which were associated with reduced male fertility in reproduction studies [24]. At the 'no observed adverse effect level' (NOAEL) in dogs, defined as the highest drug exposure level that produced no adverse effects, the exposure was twofold above the exposure seen in humans taking filgotinib 200 mg once daily [24], the highest clinical dose evaluated in phase 3 studies [25, 27].

Semen parameters are among very few clinical parameters that can be used to reliably

monitor testicular function, along with serum testosterone and gonadotropin (e.g. follicle stimulating hormone [FSH] and luteinizing hormone [LH]) measurements [28]. In evaluating parameters related to testicular function, it is important to account for the known latency period of several months (reflecting the approximately 13-week time course of one spermatogenic cycle) between testicular injury and detection of its effect on semen parameters [29]. In addition, there is a limited ability to interpret changes from baseline in semen parameters in terms of their effects on fertility [28]. Other measures of relevance to findings with filgotinib in animal studies, such as pregnancy rates to assess fertility, or evaluations of testicular histology, are neither practical nor feasible for use in clinical trials [28]. Semen parameters are thus the main outcome measures recommended for trials evaluating testicular toxicity and/or fertility in men, and US Food and Drug Administration (FDA) guidance is provided on the conduct of such trials [28].

On the basis of the aforementioned findings with filgotinib in animals, and upon consultation with global regulatory authorities, two international, randomized, double-blind, placebo-controlled, parallel-group phase 2 studies (MANTA [EudraCT number 2017-000402-38; ClinicalTrials.gov identifier NCT03201445] and MANTA-RAY [EudraCT number 2018-003933-14; ClinicalTrials.gov identifier NCT03926195]) were designed to evaluate the impact (if any) of filgotinib 200 mg on semen parameters in participants with IBD and rheumatic diseases, respectively. These trials were designed so that data for the week-13 primary endpoint could be pooled for analysis, thus encompassing a wide range of participants with inflammatory diseases.

Here we present the rationale and methodology for the MANTA and MANTA-RAY studies, with the goal of providing appropriate context for interpretation of their forthcoming results, as well as the practical insights gained during the design of these trials.

RATIONALE FOR PATIENT SELECTION

Recent (finalized October 2018) FDA guidelines for the evaluation of testicular toxicity during drug development recommend that, to the extent feasible, participants should be representative of the population(s) for whom the drug is intended [28]. This guidance is especially pertinent to chronic inflammatory conditions, which are known to affect the male reproductive system [6, 7, 14]. Therefore, effects of filgotinib on semen parameters in healthy individuals may not sufficiently reflect those in patients with inflammatory diseases. The MANTA and MANTA-RAY studies therefore included participants with IBD and rheumatic diseases, respectively, with inadequate response (or intolerance) to previous inflammatory-disease-specific treatment.

General Eligibility Criteria

Both the MANTA and MANTA-RAY studies included men aged 21–65 years with active inflammatory disease (Table 1). Full inclusion and exclusion criteria for these studies are provided in Tables S1–S4 in the electronic supplementary material. In MANTA, participants were required to have documented UC (minimum extent of 15 cm from the anal verge) or CD lasting at least 4 months, and to meet the criteria for moderately to severely active IBD at (or in the 90 days before) screening. In MANTA-RAY, participants had to meet specific diagnostic criteria for RA, psoriatic arthritis (PsA), ankylosing spondylitis (AS) or non-radiographic axial spondyloarthritis (nrAxSpA), for at least 12 weeks before screening, and criteria for active disease during the screening period (Table 1).

Individuals were excluded from MANTA and MANTA-RAY if they had previously documented problems with male reproductive health (e.g. primary or secondary hypogonadism) or a previous diagnosis of reduced fertility. Choice of eligibility criteria for baseline semen parameters was primarily based on FDA guidance, which recommends values equal to or exceeding the generally accepted 5th percentile

of World Health Organization (WHO) 2010 reference values [30]. The eligibility criteria used in the current study were semen volume ≥ 1.5 mL, total sperm per ejaculate ≥ 39 million, sperm concentration ≥ 15 million/mL, sperm total motility $\geq 40\%$ and normal sperm morphology $\geq 30\%$. These were based on the mean of two separate semen samples taken during the 45-day screening period (further procedural details are provided later). Motility assessments were based on the proportion of sperm with any sort of movement (total motility), rather than the proportion of sperm with forward movement (progressive motility).

The aforementioned reference values are based on data from ‘fertile’ men (defined as those with partners who had a time to pregnancy of ≤ 12 months) and do not represent a ‘normal’ reference range per se, but were developed as a tool for use in conjunction with clinical data to contribute to clinical evaluations of semen quality and prospects for fertility [30]. Exceptions to the WHO 2010 criteria in the MANTA and MANTA-RAY trials were total sperm motility, with a 40% cut-off (vs progressive motility, with a 32% cut-off in the WHO 2010 criteria [31]), and normal morphology, with a 30% cut-off (vs 4% in the WHO 2010 criteria [31]), which were based instead on WHO 1992 criteria [32, 33]. These exceptions reflect the fact that development of WHO 1992 criteria (vs WHO 2010 criteria) was based more on studies of testicular toxicity/sperm safety (vs studies in fertile men) [33], making them potentially more suitable in this instance.

Exclusion of Potentially Sperm-Modifying Medications

While data on the effects of anti-inflammatory agents on semen parameters are limited, sulfasalazine has been shown to have a clear negative effect on semen parameters [16]. Conversely, studies investigating exposure to tumour necrosis factor inhibitors in patients with AS have indicated possible improvements in semen parameters with these drugs [16]. Consistent with FDA guidance, use of these (and other) potential sperm-modifying agents

Table 1 Key inclusion criteria for the MANTA and MANTA-RAY studies (see Tables S1–S4 in the electronic supplementary material for full inclusion and exclusion criteria)

MANTA	MANTA-RAY
Men aged 21–65 years (inclusive) on the day of signing informed consent	Same as MANTA
Semen parameters	
Mean of two separate semen samples collected during the screening period meeting the following minimum criteria:	
Semen volume \geq 1.5 mL (WHO 2010 criteria)	
Total sperm ejaculate \geq 39 million sperm (WHO 2010 criteria)	
Sperm concentration \geq 15 million/mL (WHO 2010 criteria)	
Sperm total motility \geq 40% (WHO 1992 criteria)	
Normal sperm morphology \geq 30% (WHO 1992 criteria)	
IBD criteria	Rheumatic disease criteria
Documented diagnosis of UC (minimum disease extent of 15 cm from the anal verge) or CD of \geq 4 months' duration. Documentation must include endoscopic and histopathological documentation of either UC or CD, as follows:	Diagnosis of RA, PsA, AS or nrAxSpA for \geq 12 weeks before screening, based on the following specific classifications:
UC	RA
Medical record documentation of, or an endoscopy report dated \geq 4 months before randomization which shows features consistent with, UC as determined by the procedure-performing physician, AND	ACR/EULAR 2010 criteria
Medical record documentation of, or a histopathology report indicating feature consistent with, UC as determined by the pathologist	PsA
CD	CASPAR criteria
Medical record documentation of, or an ileocolonoscopy (full colonoscopy with intubation of terminal ileum) reported dated \geq 4 months before randomization, which shows features consistent with, CD as determined by the procedure-performing physician, AND	AS or nrAxSpA
Medical record documentation of, or a histopathology report indicating features consistent with, CD as determined by the pathologist	ASAS criteria
At (or in the 90 days before) screening, patients had to meet the following criteria for moderately to severely active disease:	At screening, patients had to meet the following criteria:
UC	RA
MCS \geq 6, PhGADA of 2 or 3 and endoscopic subscore \geq 2	CDAI-RA score $>$ 10
CD	PsA
CDAI total score \geq 220, AND	DAPSA score $>$ 14
Evidence of active inflammation, with a total SES-CD score of \geq 6 OR if disease is limited to the ileum and/or right colon, a combined SES-CD score \geq 4 in these two segments	AS or nrAxSpA
	BASDAI score \geq 4
	hsCRP $>$ 0.3 mg/dL OR sacroiliitis according to the modified New York criteria OR a history of active inflammation on magnetic resonance imaging consistent with sacroiliitis within 2 years of screening

Table 1 continued

MANTA	MANTA-RAY
Previous treatment failure or intolerance	Previous treatment failure or intolerance
Previously inadequate clinical response, loss of response or intolerance to at least one of the following five classes of agent:	RA
Corticosteroids	Inadequate response or intolerance to ≥ 12 weeks of csDMARD or bDMARD therapy for RA
Azathioprine, 6-mercaptopurine or methotrexate	PsA
TNF α antagonists (infliximab, adalimumab, golimumab [UC only] or certolizumab [CD only])	Inadequate response or intolerance to ≥ 12 weeks of csDMARD or bDMARD therapy for PsA
Vedolizumab	AS or nrAxSpA
Ustekinumab (CD only)	Inadequate response or intolerance to ≥ 12 weeks of csDMARD or bDMARD therapy for spondyloarthritis OR
	Inadequate response or intolerance to at least two NSAIDs, which may include COX-2 inhibitors prescribed for a period of ≥ 4 weeks

ACR American College of Rheumatology, *AS* ankylosing spondylitis, *ASAS* Assessment of SpondyloArthritis international Society, *BASDAI* Bath Ankylosing Spondylitis Disease Activity Index, *bDMARD* biologic disease-modifying anti-rheumatic drug, *CASPAR* Classification Criteria for Psoriatic Arthritis, *CD* Crohn’s disease, *CDAI* Crohn’s Disease Activity Index, *CDAI-RA* Clinical Disease Activity Index for Rheumatoid Arthritis, *COX-2* cyclooxygenase-2, *csDMARD* conventional synthetic disease-modifying anti-rheumatic drug, *DAPSA* Disease Activity in Psoriatic Arthritis, *EULAR* European League Against Rheumatism, *hsCRP* high-sensitivity C-reactive protein, *IBD* inflammatory bowel disease, *MCS* Mayo Clinic Score, *nrAxSpA* non-radiographic axial spondyloarthritis, *NSAID* non-steroidal anti-inflammatory drug, *PbGADA* Physician Global Assessment of Disease Activity, *PsA* psoriatic arthritis, *RA* rheumatoid arthritis, *SES-CD* Simple Endoscopic Activity Score in Crohn’s Disease, *TNF α* tumour necrosis factor α , *UC* ulcerative colitis

was prohibited before and during the MANTA and MANTA-RAY studies owing to their potential confounding effects on the trial results [28].

Other Medication Restrictions

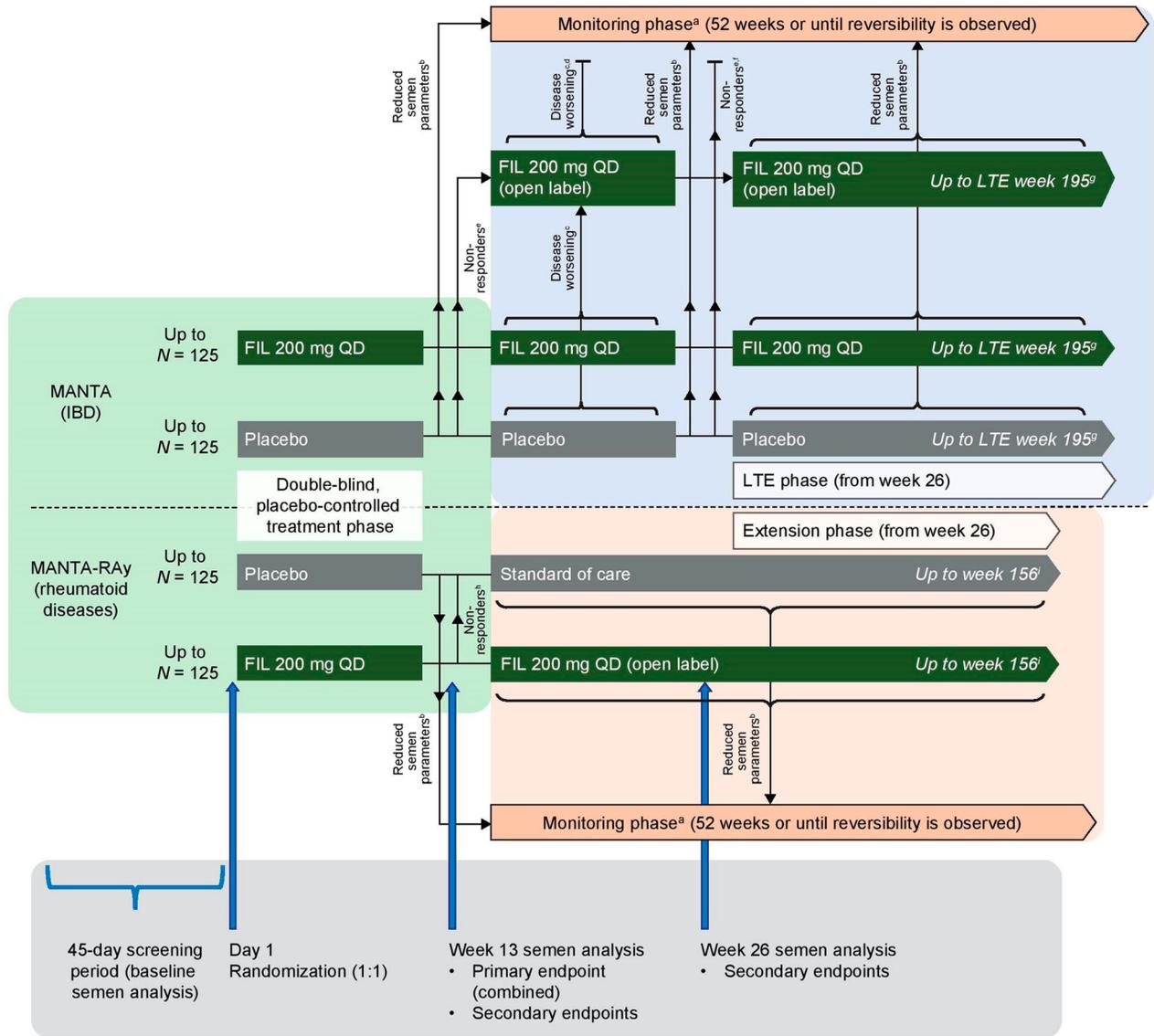
Potential sperm-modifying effects have also been implicated for other medications, but data are so far inconclusive [16]. Exclusion of all such drugs in the MANTA trial, on the basis of ‘possible’ sperm-modifying effects, would have been untenable from a recruitment perspective. In the MANTA trial, concomitant use of oral 5-aminosalicylic acid compounds, azathioprine, 6-mercaptopurine, methotrexate and corticosteroids (dose equivalent to ≤ 20 mg/day of prednisone) was therefore allowed, provided the prescribed dose was stable for at least 4 weeks before randomization and up to, at minimum, the time of primary endpoint assessment (methotrexate required a stable dose for 26 weeks). In the MANTA-RAY trial,

concomitant use of methotrexate (≤ 25 mg/week), leflunomide, hydroxychloroquine, chloroquine and apremilast was allowed, also with the requirement of stable dosing at least 4 weeks before randomization and up to, at a minimum, primary endpoint assessment.

Effective IBD and rheumatic disease therapies were not discontinued for the purposes of inclusion in these trials. The full lists of medications that were restricted during these trials are presented in Tables S5 and S6 in the electronic supplementary material.

Logistical Considerations

IBD occurs during the peak reproductive years. It is associated with a higher frequency, versus the general population, of both reproductive (e.g. reduced sperm quality [6], increased risk of erectile dysfunction [7]) and psychological issues (e.g. reduced libido [7], impaired body image [8], anxiety and depression [9]). All of



◀ **Fig. 1** MANTA and MANTA-RAY study designs. ^aStudy drug was discontinued upon entry into the monitoring phase (standard of care was initiated [or continued] in MANTA-RAY). Reversibility for a participant was defined as all semen parameters^b that qualified the participant for entry into the monitoring phase returning to > 50% of baseline. ^bReduced semen parameters were defined as a $\geq 50\%$ decrease in sperm concentration, and/or motility, and/or morphology compared with baseline. ^cDisease worsening for UC in MANTA was defined as an increase of ≥ 3 points in pMCS (to a score of ≥ 5) from week 13, on two consecutive visits, or an increase to a score of 9 (from a week 13 score of > 6) on two consecutive visits; disease worsening for CD in MANTA was defined as an increase of ≥ 100 points in CDAI score from the week 13 visit on two consecutive visits, and a score of ≥ 220 on two consecutive visits. ^dIn MANTA, participants who had disease worsening^c between weeks 13 and 26 while on open-label filgotinib discontinued the study and completed an ET visit, and had a safety visit 30 days after the last dose of study drug. ^eResponse for UC in MANTA was defined as a reduction of ≥ 2 points in pMCS compared with the baseline visit. Response for CD in MANTA was defined as a reduction of ≥ 100 points in total CDAI score compared with the baseline visit; in patients with a total CDAI score of ≥ 220 (but ≤ 250) at the baseline visit, response was defined as a CDAI score of < 150. ^fIn MANTA, non-responders^e at week 26 discontinued the study and had only a safety visit 30 days after the last dose of study drug. ^gIn MANTA, participants who had disease worsening^c during the LTE (receiving open-label filgotinib or blinded study drug [filgotinib or placebo]) discontinued and completed an ET visit, followed by a safety follow-up visit 30 days after last study drug dose; participants receiving open-label filgotinib or blinded study drug who had reduced semen parameters^b (assessed every 13 weeks) during the LTE entered the monitoring phase. ^hResponse for rheumatoid diseases in MANTA-RAY was defined as an improvement of $\geq 20\%$ in PhGADA score compared with day 1. ⁱIn MANTA-RAY, participants who were receiving open-label filgotinib or standard of care and required prohibited concomitant treatment discontinued the study and completed an ET visit. Safety visits occurred 30 days after the last study drug dose. *CD* Crohn's disease, *CDAI* Crohn's Disease Activity Index, *ET* early termination, *FIL* orally administered filgotinib, *IBD* inflammatory bowel disease, *LTE* long-term extension, *PhGADA* Physician's Global Assessment of Disease Activity, *pMCS* partial Mayo Clinic Score, *QD* once daily, *UC* ulcerative colitis

these factors may combine to affect sexual and reproductive health. Nevertheless, fecundity in patients with IBD is not necessarily reduced versus the general population [10], highlighting that overall fertility is a complex product of both male and female sexual and reproductive health [10]. Voluntary childlessness, however, is more common in patients with IBD than in the general population, possibly owing to heightened concern about adverse reproductive outcomes that, although legitimate, may be disproportionate to the available evidence [10, 11].

Rheumatic disease can also affect sexual function and reproduction in men, with sexual dysfunction reported in a higher proportion of men with rheumatological conditions compared with healthy controls [12–14]. In addition, the recently conducted multicentre, cross-sectional iFAME-Fertility study found that participants diagnosed with inflammatory arthritis before and during their peak reproductive age had a lower fertility rate, a higher childlessness rate and an increased frequency of fertility problems, compared with those who were diagnosed after their peak reproductive age [15]. Importantly, some treatments for RA (and IBD) may also affect sexual and reproductive health [16].

In light of the aforementioned considerations, it was considered possible that the highly restrictive (IBD and reproductive) eligibility criteria, and restriction of some widely used medications, particularly sulfasalazine, had the potential to substantially limit the recruitment pool. These factors contributed to the decision to use an international, multicentre study design, to ensure that enrolment rates were sufficient to complete the trials in a timely manner.

It was also anticipated during the design of the MANTA and MANTA-RAY trials that patients with inflammatory diseases, who often have negative perceptions concerning fertility, may be hesitant to participate in trials evaluating semen parameters, which are often mistakenly equated to measures of fertility and sexual performance [11, 34]. With regard to the latter point, materials were provided to help normalize and simplify the requirement for a semen

sample, to educate study staff and patients on men's reproductive health to enable accurate and comfortable conversations, and to help study staff and patients understand the objectives and value of the MANTA and MANTA-RAY studies.

STUDY DESIGN RATIONALE

Both MANTA and MANTA-RAY were conducted in accordance with the principles of the Declaration of Helsinki, International Council for Harmonisation guidelines, or with the laws and regulations of the country in which the research was conducted, whichever afforded the greatest protection to study participants. Trial protocols and subsequent amendments were approved by the relevant institutional review board at each participating site. All patients provided written informed consent before enrolment. Both studies included a 45-day screening period for assessment of baseline parameters, followed by randomization (1:1) to receive either filgotinib 200 mg or placebo orally, once daily for 13 weeks (allowing pooled analysis of the primary endpoint), after which disease-specific considerations required some deviations in the designs of these trials (Fig. 1). In both trials, randomization was stratified according to disease type (UC vs CD in MANTA; RA vs spondyloarthritis [PsA, AS or nrAxSpA] in MANTA-RAY), baseline sperm concentration (15–25 million/mL vs > 25–50 million/mL vs > 50 million/mL) and methotrexate use (yes vs no).

FDA guidelines recommend that semen analyses for the assessment of drugs intended for chronic use should be conducted at baseline, and at the end of each of two consecutive 13-week treatment periods (equivalent to two full spermatogenic cycles), and at least 13 weeks after study drug discontinuation [28]. MANTA and MANTA-RAY were designed to accommodate these requirements, up to the point where they conflicted with ethical considerations in this population, namely the need to switch to an alternative effective therapy if the subject is not responding to treatment after 13 weeks, especially with respect to the prevention of

irreversible, inflammatory disease-related damage.

The double-blind treatment phase of the MANTA study consisted of two consecutive parts (A and B), each comprising a 13-week treatment period (Fig. 1). Stable dosing of permitted concomitant medications was allowed during part A, with dosing changes allowed thereafter (within specified limits). Participants who had a pre-specified decrease in one or more semen parameters from baseline ($\geq 50\%$ decrease in sperm concentration, and/or total motility and/or normal morphology) at the week 13 study visit, or any semen parameter assessment visit thereafter, discontinued study drug and entered a monitoring phase, during which they were offered investigator-selected standard of care therapy as per protocol, as long as it did not have known effects on semen parameters. During the monitoring phase, the reduced semen parameter (or parameters) that qualified the patient for entry into the monitoring phase was evaluated every 13 weeks for up to 52 weeks, or until reversibility was observed, defined as the return of that semen parameter to greater than 50% of the baseline value. Participants who did not have a pre-specified decrease in one or more semen parameters from baseline at the end of part A (week 13), and who were also disease responders, continued blinded study treatment in part B (weeks 13–26), while those who were disease non-responders were switched to open-label filgotinib 200 mg. Participants with disease worsening in part B (weeks 13–26) while on blinded study treatment were switched to open-label filgotinib 200 mg, and those with disease worsening while on open-label filgotinib 200 mg were discontinued. Responders at week 26 (blinded or open-label filgotinib) who had not had a pre-specified decrease from baseline in one or more semen parameters continued in the long-term extension (LTE). Treatment in the LTE could be continued for up to 195 weeks if participants continued to meet the LTE entry criteria (assessed every 13 weeks).

The MANTA-RAY study was designed to be similar to MANTA up to week 13, when the primary endpoint (proportion of participants with a $\geq 50\%$ decrease from baseline in sperm

concentration) was assessed—allowing for pooling of data from the two studies at the week-13 timepoint (Fig. 1). Beyond this time point, the key difference (vs MANTA) was that the MANTA-RAY study no longer provided a placebo comparator, owing to the potential for irreversible tissue damage in untreated participants with RA. Thus, at the end of the 13-week double-blind treatment phase in MANTA-RAY, all disease non-responders, and disease responders originally assigned to placebo, were switched to standard of care therapy, and disease responders originally assigned to filgotinib were switched to open-label filgotinib 200 mg. Participants in this ‘extension phase’ could continue treatment up to week 156. As with the MANTA study, participants with pre-specified decreases in semen parameters at any semen assessment visit discontinued study drug and entered a monitoring phase (Fig. 1).

RATIONALE FOR CHOSEN ASSESSMENTS

A summary of study procedures and assessments performed in the 13-week double-blind treatment phases of the MANTA (part A) and MANTA-RAY studies is provided in Table 2. Procedures and assessments for the MANTA (part B) double-blind phase, MANTA open-label, monitoring and LTE phases, and MANTA-RAY extension and monitoring phases are provided in Tables S7–S10 in the electronic supplementary material.

Inflammatory Disease Measures

Physician- and patient-reported outcome measures, as well as lower gastrointestinal investigations in participants with IBD, were performed at screening to determine disease activity status (Tables 1 and 2). In MANTA, partial Mayo Clinic Scores and Crohn’s Disease Activity Index scores were assessed in participants with UC and CD, respectively, to evaluate disease response to treatment (Table 1 and Fig. 1). In MANTA-RAY, Physician Global Assessment of Disease Activity scores were

assessed at screening, day 1 and week 13 to determine participants’ rheumatic disease activity status (Table 1 and Fig. 1). These are well-established inflammatory disease measures that have been used in previous clinical trials of filgotinib [25, 27, 35].

Semen Analysis Methods and Logistics

There is substantial natural variation in an individual’s sperm concentrations and other semen parameters over time [36], independent of known modifiers such as fever [37] and certain anti-inflammatory medications [16]. Thus, sperm concentrations may occasionally fall below the 5th percentile of WHO 2010 reference values, even in individuals who might typically produce values above this threshold [36]. The collection of two separate semen samples for each evaluation of semen parameters, and the requirement for minimum and maximum abstinence periods between ejaculations, substantially increased the logistical complexity of the MANTA and MANTA-RAY studies, with study sites (or associated fertility clinics) and technicians all required to be available within 1 h of a sample being provided. However, this requirement was necessary to minimize intra-individual variation in semen parameters, optimize screening accuracy and ensure differences between treatment groups could be detected if present.

In both MANTA and MANTA-RAY, semen samples were collected at screening and then every 13 weeks throughout each phase of each study, as well as at early termination visits. For every assessment stage where semen parameters were measured, two separate semen samples were collected within 14 days of each other. Each semen sample had to be collected within an ejaculation-free period of at least 48 h and no greater than 7 days. In select instances where the semen sample was found to be non-assessable or invalid (e.g. collection without adherence to ejaculation-free periods; intercurrent illness or dehydration at time of collection; incomplete capture of semen sample; and/or sample processing or semen analysis that deviated from standardized procedures), a third

Table 2 MANTA (part A) and MANTA-RAy procedures and assessments during the double-blind treatment phase

MANTA and MANTA-RAy assessments	Screening (45 days)	Day 1	Week 2 (± 5)	Week 4 (± 5)	Week 8 (± 5)	Week 13 ^a (+ 5)	ET ^b	Safety follow-up ^c
Informed consent	X							
Randomization		X						
Demographics and medical history	X							
Inclusion/exclusion criteria review	X	X						
Complete physical exam	X						X ^d	X ^d
Symptom-directed physical examination, as needed		X	X	X	X	X	X ^c	X ^c
Vital signs and weight	X	X	X	X	X	X	X	X
Height	X							
12-lead ECG	X					X ^d	X ^d	
TB test	X							
Chest X-ray	X							
Urinalysis	X	X				X	X	X ^d
Urine drug screen	X							
Haematology and serum chemistry	X	X	X	X	X	X	X	X
Lipid profile (fasting)		X				X		
TSH and HbA1c (endocrine)	X							
FSH, LH, inhibin B, total testosterone (sex hormones)	X	X		X		X	X	X
Blood samples for pharmacokinetics			X	X		X		
Concomitant medications	X	X	X	X	X	X	X	X
Adverse events	X	X	X	X	X	X	X	X
Semen collection (two samples) ^f	X					X	X ^g	
Date and time of most recent ejaculation ^h	X					X	X	
HIV, hepatitis B and hepatitis C	X							
HBV DNA						X		
MANTA-specific assessments								
Flexible sigmoidoscopy	X							
Ileocolonoscopy	X							
Complete MCS (for UC only)	X							
Partial MCS (for UC only)		X	X	X	X	X	X	
SES-CD (for CD only)	X							
CDAI (for CD only)	X	X	X	X	X	X	X	
Serum immunoglobulin	X					X		
MANTA-RAy-specific assessments								
TJC28 score (for RA only)	X							
SJC28 score (for RA only)	X							
TJC68 score (for PsA only)	X							
SJC66 score (for PsA only)	X							
PtGADA score (for RA and PsA)	X							
PhGADA score	X	X				X		

Table 2 continued

MANTA and MANTA-RAY assessments	Screening (45 days)	Day 1	Week 2 (± 5)	Week 4 (± 5)	Week 8 (± 5)	Week 13 ^a (+ 5)	ET ^b	Safety follow-up ^c
PPain (for PsA only)	X							
BASDAI (for AS/nrAxSpA only)	X							
Serum hsCRP	X	X	X	X	X	X		

BASDAI Bath Ankylosing Spondyloarthritis Disease Activity Index, *CD* Crohn’s Disease, *CDAI* Crohn’s Disease Activity Index, *DNA* deoxyribose nucleic acid, *ECG* electrocardiogram, *ET* early termination, *FSH* follicle-stimulating hormone, *HbA1c* glycated haemoglobin, *HBV* hepatitis B virus, *HIV* human immunodeficiency virus, *hsCRP* high-sensitivity C-reactive protein, *LH* luteinizing hormone, *MCS* Mayo Clinic Score, *nrAxSpA* non-radiographic axial spondyloarthritis, *PrGADA* Patient Global Assessment of Disease Activity, *PPain* Patient Global Assessment of Pain Intensity, *PhGADA* Physician Global Assessment of Disease Activity, *PsA* psoriatic arthritis, *RA* rheumatoid arthritis, *SES-CD* Simple Endoscopic Score for Crohn’s Disease, *SJC28* Swollen Joint Count 28, *SJC66* Swollen Joint Count 66, *SoC* standard of care, *TB* tuberculosis, *TJC28* Tender Joint Count 28, *TJC68* Tender Joint Count 68, *TSH* thyroid-stimulating hormone, *UC* ulcerative colitis

^aWeek 13 visit must occur after 13 weeks of drug exposure. Therefore, visit window for this visit is + 5 days. The end of the double-blind treatment phase is defined as occurring when the week-13 semen results have been evaluated by the investigator. On the basis of this evaluation, subjects are assigned to enter either the monitoring phase (participants with ≥ 50% decrease in sperm concentration, and/or motility and/or morphology), open-label filgotinib treatment (MANTA study responders) or the extension phase (MANTA-RAY study: open-label filgotinib treatment [responders] or SoC [non-responders])

^bAny time a subject discontinued participation in the study before week 26 in MANTA, or before week 13 in MANTA-RAY, an ET visit was required. Subjects who were responders at week 26 in MANTA, or at week 13 in MANTA-RAY, and who chose not to continue in the open-label filgotinib phase (MANTA) or extension phase (MANTA-RAY), respectively, were also required to complete ET visit assessments

^cAll subjects had a 30-day safety follow-up visit after discontinuing study drug, even if they entered the monitoring phase

^dMANTA-RAY study only

^eMANTA study only

^fThe screening semen sample collection coincided with the day 1 (baseline) visit where possible. Semen samples were collected on the visit day and/or as soon as possible after the visit day

^gSemen collection was completed at the ET visit only if previous semen samples were not collected within 2 weeks of ET

^hAsked before each semen collection

semen sample was to be collected within 14 days of the date the previous sample was collected.

FDA guidelines recommend that the collection and handling of semen samples should be standardized for all sites, and that a single central laboratory should process and analyse all semen samples for the purposes of consistency and quality assurance [28]. However, as described, an international, multicentre approach was employed in both the MANTA and MANTA-RAY studies to ensure timely completion of the study to address potential recruitment difficulties. The requirement for fresh semen samples precluded centralized laboratory assessments of semen parameters (except for sperm morphology), especially for motility which must be measured within 1 h of sample collection.

To ensure consistency across study sites, stringent standardization processes (with central oversight by staff at Tulane University, New Orleans, LA, USA) were applied. Suitable candidate technicians for training were selected

through live instruction and written tests of proficiency in the assessment of semen parameters. Technicians who successfully passed the training programme were issued a Certificate of Completion from Tulane University, followed by immediate post-training evaluations of technicians’ proficiency at their ‘home’ laboratory, and follow-up assessments to ‘requalify’ technicians every 6 months. Timely refresher calls for troubleshooting of any issues that arose, and quality control (QC) review at Tulane University of all semen analysis data to verify completion of all fields as well as technicians’ calculations, were also used to maintain consistency. In addition, staining and reading of sperm morphology smears (prepared at the study site or fertility clinic using fresh semen samples) were performed at Tulane University. Study sites and fertility clinics also had the option to use qualified ‘travelling technicians’ provided by a vendor, as either a primary option for semen analysis or as a backup option.

Inclusion of Sex Hormone Assessments

Hormone measures are routinely included in the clinic in addition to assessments of semen parameters, to obtain a more granular picture of testicular health, and (as already described) are among the few clinical parameters that can be used to reliably monitor testicular function [28]. Although no firm evidence of hormone abnormality indicative of testicular injury was observed with filgotinib treatment in the pre-clinical animal studies, blood samples were nevertheless collected for the measurement of sex hormone levels (FSH, LH, inhibin B and total testosterone) in the MANTA and MANTA-RAY studies (Table 2 and Tables S7–S10 in the electronic supplementary material).

Routine Safety Assessments

Safety evaluations (adverse events, haematology and serum chemistry, vital signs, body weight and symptom-directed medical examinations) and assessments of concomitant medication use occurred at every study visit. Electrocardiograms and further laboratory measures (urinalysis, lipids) were performed at predefined time points throughout each study phase (Table 2 and Tables S7–S10 in the electronic supplementary material).

Additional Safety Assessments

In addition to routine safety evaluations, an independent committee periodically reviewed and adjudicated all potential major adverse cardiovascular events (cardiovascular death, myocardial infarction and stroke) and thromboembolic events (arterial thrombosis and venous thrombosis) in a blinded manner. It is noteworthy that identification of any such signals by the independent committee would lead to early interruption and unblinding of the MANTA and MANTA-RAY trials.

RATIONALE FOR SAMPLE SIZE, ENDPOINTS AND THEIR ANALYSIS

Sperm concentration is considered the most reliable quantifiable parameter for providing information about effects on testicular function [28], and may be the most important stand-alone indicator of such effects [33]. However, there are no well-established differences that would be clinically relevant between treatment arms in terms of the proportions of participants with a reduction of at least 50% from baseline in sperm concentration (primary endpoint) [28]. Analyses for MANTA and MANTA-RAY were thus not powered to detect predefined differences between treatment arms for the primary endpoint. In recognition of the fact that no single semen parameter is adequate for determining drug effects on testicular function, and in line with FDA guidance, an array of additional secondary and exploratory endpoints were also evaluated, with conclusions regarding the effects of filgotinib on testicular function to be determined using the totality of the data across all endpoints [28].

Primary Endpoint

The primary endpoint in the MANTA and MANTA-RAY trials was the proportion of participants, pooled across both trials, with a $\geq 50\%$ decrease from baseline in sperm concentration at week 13 (equivalent to one full spermatogenic cycle). This is a surrogate biomarker of testicular function that has been used in numerous trials and is the FDA-recommended primary endpoint [28, 38].

It is worth noting that a 50% reduction from baseline in sperm concentration, which is within the range of natural variation [36], will not necessarily result in absolute values that are below 15 million/mL (i.e. the 5th percentile of WHO 2010 reference values [31]). For example, an individual with a baseline sperm count of 50 million/mL will still have a sperm count of 25 million/mL after a 50% reduction, thus still exceeding this threshold value. Similarly, reductions in sperm count of less than 50% may lead to values below 15 million/mL in some

individuals. In addition, high intra-individual variation in sperm concentrations may lead to regression towards the mean, that is, individuals with high ‘outlier’ values at baseline are more likely to have a 50% decrease at week 13 by chance alone, owing to the tendency of measurements at future points to be closer to the mean. Negative trends in semen parameters in placebo-treated individuals in clinical trials have been partly attributed to regression towards the mean [39], an effect likely to be enhanced by initial selection of individuals with minimum semen parameter values.

Secondary Endpoints

The proportion of participants with at least a 50% decrease from baseline in sperm concentration at week 26 was assessed as a key secondary endpoint. Additional secondary endpoints relating to semen parameters, evaluated at weeks 13 and 26, included changes from baseline in sperm concentration, sperm total motility, sperm morphology, total sperm count and ejaculate volume. It should be reiterated that only the MANTA study included a placebo comparator for week 26 outcomes (Fig. 1), for reasons already described.

Exploratory Endpoints

In participants with at least a 50% decrease from baseline in sperm concentration, and/or motility and/or morphology at the week 13 or week 26 semen assessment visits in each trial, or at any of the semen assessment visits occurring every 13 weeks during the MANTA LTE or MANTA-RAY extension phases, the reversibility of these effects was evaluated every 13 weeks during a 52-week monitoring phase (Fig. 1). Additional exploratory endpoints included effects on sex hormones (FSH, LH, inhibin B and total testosterone) at weeks 13 and 26.

Evaluations of the pharmacokinetics of filgotinib and its metabolite (GS-829845) as well as the long-term safety and tolerability of filgotinib (assessed in the MANTA LTE and MANTA-RAY extension phases—Fig. 1) were also included as exploratory endpoints.

Statistical Analysis

As noted earlier, there are no well-established differences in semen parameters that would be considered clinically relevant between treatment arms [28]. This is reflected in the methods of data presentation and analysis, which were not based on a hypothesis-driven assessment of minimum clinically important differences between treatments [28].

Each trial could recruit up to 125 participants per arm. The overall goal was to have a combined total of at least 200 evaluable participants for the pooled primary endpoint analysis of week-13 data from both MANTA and MANTA-RAY. This sample size is adequate for estimating cumulative distribution curves and producing a reasonably narrow 95% confidence interval (CI) width for the difference between treatment groups in the proportion of participants with at least a 50% decrease from baseline in sperm concentration. For the primary endpoint, a cumulative distribution plot for the percentage change from baseline in sperm concentration at week 13 was constructed, with the *x*-axis displaying percentage changes from baseline (range –100% to the maximal observed increase), and the *y*-axis displaying the proportion of participants with a percentage change in sperm concentration equal to or less than the corresponding *x*-axis value. Proportions of participants experiencing at least a 50% decrease from baseline in sperm concentration were calculated for each treatment group. Differences between the filgotinib and placebo groups were calculated together with associated 95% CIs.

For secondary endpoints, continuous variables were summarized using descriptive statistics, while categorical variables were summarized by the number/percentage of participants meeting endpoint definitions. For exploratory analyses of reversibility, numbers and percentages (point and cumulative data) of participants achieving reversibility in semen parameters were displayed by semen parameter and monitoring phase visit.

DISCUSSION

Here we report the rationale and methodology for the phase 2 MANTA and MANTA-RAY studies, which were designed to elucidate the clinical relevance (if any) in man of observed effects of filgotinib on semen parameters in preclinical animal studies [24, 40].

A significant challenge for the MANTA and MANTA-RAY studies was the need to include patients with active disease who were also willing to participate in the trials. There may be a hesitancy to be involved in trials of this nature because semen parameter measures may be mistakenly viewed as equivalent to measures of fertility and sexual performance, and steps were therefore taken to allay such fears. Exclusion of individuals with reproductive health or fertility issues, and with baseline semen parameters below the 5th percentile of the WHO reference values, likely also reduced the pool of eligible individuals, and may have selected for a population that is less representative of individuals with inflammatory diseases who would receive filgotinib.

FDA criteria recommend that drugs intended for chronic use should be administered for 26 weeks (two full spermatogenic cycles), compared with 13 weeks for drugs intended for shorter-term use. Ideally, both 13-week phases should be performed under randomized, blinded conditions with stable dosing of permitted concomitant medications. However, ethical considerations—such as the need for participants to be able to switch to a more effective therapy to avoid disease-related damage—demanded greater flexibility from week 13 onwards for participants who did not respond to study treatment. This added a challenging layer of complexity to the current trials, which would also apply to any randomized trial that aims to determine the effects of a drug on semen parameters in patients with certain chronic inflammatory conditions.

The use of a multicentre approach to the MANTA and MANTA-RAY trials helped to ensure that the minimum recommended number of participants could be achieved in a timely manner despite barriers to recruitment.

However, this meant that assessments of semen parameters (except for morphology) could not be centralized. Thus, rigorous processes (e.g. stringent technician selection, initial training and follow-up training, use of travelling technicians and centralized QC methods, all with strict central oversight) were implemented to ensure that semen parameter measures were standardized across the study centres, which spanned numerous countries.

Despite being used as a surrogate marker for gonadal function, and being a key component of clinical explorations of fertility, negative effects on semen parameters do not necessarily equate to reduced fertility. As an equation involving two partners, fertility assessment should account for factors such as the overall health and fertility of the female partner, which also contribute to the chances of natural conception [41, 42]. IBD is associated with an increased frequency of several sexual health issues that may impede conception, including erectile dysfunction, reduced libido, impaired body image, and anxiety, which probably contribute to negative impacts on reproductive health [7–9]. Sexual function and reproduction can similarly be impaired in patients with rheumatic diseases [12–14]. In addition, inflammation itself may directly impair hormonal production and sperm quality in patients with inflammatory diseases [6, 14, 43–45]. It is reasonable to expect that successful treatment of inflammatory diseases may alleviate some of these factors to improve overall fertility.

Sexual health and reproductive health have distinct components, but are also inherently interlinked [46]. For example, concerns about the effects of diseases or treatments on reproductive health may have emotional and psychological consequences that affect sexual health. Conversely, diseases or treatments that affect aspects of sexual health, such as sexual function, may affect reproductive health. A recent post hoc analysis of data from the phase 3 FINCH trials investigating filgotinib for the treatment of RA reported that baseline disease activity negatively impacted sexual function in both male and female patients, and that active treatment with filgotinib or adalimumab resulted in early and sustained improvements

from baseline in sexual function [47]. The net effects of anti-inflammatory drugs on components of fertility must therefore be carefully considered in this population.

CONCLUSIONS

In summary, despite the complexities associated with the investigation of potential drug effects on semen parameters in patients with chronic inflammatory diseases, unique and robust trial designs were achieved for the MANTA and MANTA-RAY studies. This landmark trial programme, including more than 200 participants, is the first large-scale, placebo-controlled evaluation of the potential impacts of an advanced therapy for immune-mediated inflammatory conditions on semen parameters. These studies may inform, and provide a template for future evaluations and offer a more holistic view of this under-studied and challenging aspect of living with a chronic inflammatory disease.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Robin Besuyen of Galapagos BV, Leiden, Netherlands, for his contributions to the design and methodology of the MANTA and MANTA-RAY studies and his interpretation of the data. Publication coordination was provided by Jessica Naddafy-Clark of Galapagos NV.

Funding. The MANTA study was sponsored by Gilead Science Inc. and conducted in collaboration with Galapagos NV. The MANTA-RAY study was sponsored by Galapagos NV and conducted in collaboration with Gilead Science Inc. Funding for medical writing support, and for the journal's Rapid Service and Open Access Fees, was provided by Galapagos NV.

Medical Writing Assistance. The authors thank Michael Molloy-Bland, PhD, of Oxford PharmaGenesis for providing medical writing support, funded by Galapagos NV, Mechelen, Belgium, in accordance with Good Publication Practice guidelines.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Author Contributions. WH, SS and DV contributed to the concept or design of the work and endorsed the trial design. RD, WR and TER were members of the study advisory board. DV, TW, CT, GD and AG contributed to the concept or design of the work. LG and SA contributed to the acquisition or analysis of the data. All authors contributed to the interpretation of the data and critically reviewed the manuscript for important intellectual content, and approved the final version of the manuscript for submission.

Disclosures. Wayne Hellstrom has been a consultant or advisor for Boston Scientific, Coloplast and Endo; an investigator for Coloplast and Endo; a lecturer for Endo; on advisory boards for Gilead, Maximus and Promescent; and is a board member, officer and trustee for Theralogix. Radboud J. E. M. Dolhain has received unrestricted research grants from Dutch Arthritis Association, UCB Pharma and Galapagos; consulting fees from Galapagos; and speaking fees from UCB Pharma, Roche, Abbvie, Genzyme, Novartis and Lilly. Timothy E. Ritter has received personal fees from AbbVie, Arena Pharmaceuticals, Ferring Pharmaceuticals, Genentech, Gilead, Gossamer, Intercept, Janssen, Eli Lilly, Pfizer, Prometheus, and Takeda. Dirk Vanderschueren received consultancy fees from Galapagos for the design and set-up of the studies. Suresh Sikka declares no competing interests. Walter Reinisch has served as a speaker for Abbvie, Aesca, Aptalis, Astellas, Centocor, Celltrion, Danone Austria, Elan, Falk Pharma GmbH, Ferring, Immundiagnostik, Mitsubishi Tanabe Pharma Corporation, MSD, Otsuka, PDL, Pharmacosmos, PLS Education, Schering-Plough, Shire, Takeda, Therakos, Vifor, Yakult; as a consultant for Abbvie, Aesca, Algernon, Amgen, AM Pharma, AMT, AOP Orphan, Arena Pharmaceuticals, Astellas, AstraZeneca, Avaxia, Roland Berger GmbH,

Bioclinica, Biogen IDEC, Boehringer-Ingelheim, Bristol-Myers Squibb, Cellerix, Chemocentryx, Celgene, Centocor, Celltrion, Covance, Danone Austria, DSM, Elan, Eli Lilly, Ernest & Young, Falk Pharma GmbH, Ferring, Galapagos, Gatehouse Bio Inc., Genentech, Gilead, Grünenthal, ICON, Index Pharma, Inova, Intrinsic Imaging, Janssen, Johnson & Johnson, Kyowa Hakko Kirin Pharma, Lipid Therapeutics, LivaNova, Mallinckrodt, Medahead, MedImmune, Millennium, Mitsubishi Tanabe Pharma Corporation, MSD, Nash Pharmaceuticals, Nestle, Nippon Kayaku, Novartis, Ocera, OMass, Otsuka, Parxel, PDL, Periconsulting, Pharmacosmos, Philip Morris Institute, Pfizer, Procter & Gamble, Prometheus, Protagonist, Provention, Quell Therapeutics, Robarts Clinical Trial, Sandoz, Schering-Plough, Second Genome, Seres Therapeutics, Setpointmedical, Sigmoid, Sublimity, Takeda, Therakos, Theravance, Tigenix, UCB, Vifor, Zealand, Zyngenia, and 4SC; as an advisory board member for Abbvie, Aesca, Amgen, AM Pharma, Astellas, AstraZeneca, Avaxia, Biogen IDEC, Boehringer-Ingelheim, Bristol-Myers Squibb, Cellerix, Chemocentryx, Celgene, Centocor, Celltrion, Danone Austria, DSM, Elan, Ferring, Galapagos, Genentech, Grünenthal, Inova, Janssen, Johnson & Johnson, Kyowa Hakko Kirin Pharma, Lipid Therapeutics, MedImmune, Millennium, Mitsubishi Tanabe Pharma Corporation, MSD, Nestle, Novartis, Ocera, Otsuka, PDL, Pharmacosmos, Pfizer, Procter & Gamble, Prometheus, Sandoz, Schering-Plough, Second Genome, Setpointmedical, Takeda, Therakos, Tigenix, UCB, Zealand, Zyngenia, and 4SC; and has received research funding from Abbvie, Aesca, Centocor, Falk Pharma GmbH, Immundiagnostik, Janssen, MSD, Sandoz, Takeda. Tim Watkins, Sarah Arterburn, Angi Gillen and Caroline Tonussi are employees and shareholders of Gilead Sciences, Inc. Goele Dekkers, Leen Gilles, Alessandra Oortwijn, Katrien Van Beneden and Dick de Vries are employees of Galapagos.

Compliance with Ethics Guidelines. The MANTA and MANTA-RAY studies were conducted in accordance with the principles of the Declaration of Helsinki, International Council for Harmonisation guidelines, or with the laws

and regulations of the country in which the research was conducted, whichever afforded the greatest protection to study participants. Trial protocols and subsequent amendments were approved by the relevant institutional review board at each participating site. All patients provided written informed consent before enrolment.

Data Availability. Data sharing is not applicable to this article as no data sets were generated or analysed for the information presented in this article.

Open Access. This article is licensed under a Creative Commons Attribution-Non-Commercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Alatab S, Sepanlou SG, Ikuta K, et al. The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *Lancet Gastroenterol Hepatol.* 2020;5(1):17–30.
2. Safiri S, Kolahi AA, Hoy D, et al. Global, regional and national burden of rheumatoid arthritis 1990–2017: a systematic analysis of the global burden of disease study 2017. *Ann Rheum Dis.* 2019;78(11):1463–71.

3. Furman D, Campisi J, Verdin E, et al. Chronic inflammation in the etiology of disease across the life span. *Nat Med*. 2019;25:1822–32.
4. Baumgart DC, Sandborn WJ. Crohn's disease. *Lancet*. 2012;380(9853):1590–605.
5. Ungaro R, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel JF. Ulcerative colitis. *Lancet*. 2017;389(10080):1756–70.
6. Valer P, Algaba A, Santos D, et al. Evaluation of the quality of semen and sexual function in men with inflammatory bowel disease. *Inflamm Bowel Dis*. 2017;23(7):1144–53.
7. Riviere P, Zallot C, Desobry P, et al. Frequency of and factors associated with sexual dysfunction in patients with inflammatory bowel disease. *J Crohns Colitis*. 2017;11(11):1347–52.
8. de Rooy EC, Toner BB, Maunder RG, et al. Concerns of patients with inflammatory bowel disease: results from a clinical population. *Am J Gastroenterol*. 2001;96(6):1816–21.
9. Bel LG, Vollebregt AM, Van der Meulen-de Jong AE, et al. Sexual dysfunctions in men and women with inflammatory bowel disease: the influence of IBD-related clinical factors and depression on sexual function. *J Sex Med*. 2015;12(7):1557–67.
10. Manosa M, Navarro-Llavat M, Marin L, et al. Fecundity, pregnancy outcomes, and breastfeeding in patients with inflammatory bowel disease: a large cohort survey. *Scand J Gastroenterol*. 2013;48(4):427–32.
11. Selinger CP, Eaden J, Selby W, et al. Inflammatory bowel disease and pregnancy: lack of knowledge is associated with negative views. *J Crohns Colitis*. 2013;7(6):e206–13.
12. Tiseo BC, Cocuzza M, Bonfa E, Srougi M, Silva CA. Male fertility potential alteration in rheumatic diseases: a systematic review. *Int Braz J Urol*. 2016;42(1):11–21.
13. Tristano AG. Impact of rheumatoid arthritis on sexual function. *World J Orthop*. 2014;5(2):107–11.
14. Perez-Garcia LF, Te Winkel B, Carrizales JP, et al. Sexual function and reproduction can be impaired in men with rheumatic diseases: a systematic review. *Semin Arthritis Rheum*. 2020;50(3):557–73.
15. Perez-Garcia LF, Roder E, Goekoop RJ, et al. Impaired fertility in men diagnosed with inflammatory arthritis: results of a large multicentre study (iFAME-Fertility). *Ann Rheum Dis*. 2021;80(12):1545–52.
16. Perez-Garcia LF, Dolhain RJEM, Vorstenbosch S, et al. The effect of paternal exposure to immunosuppressive drugs on sexual function, reproductive hormones, fertility, pregnancy and offspring outcomes: a systematic review. *Hum Reprod Update*. 2020;26(6):961–1001.
17. Danese S, Allez M, van Bodegraven AA, et al. Unmet medical needs in ulcerative colitis: an expert group consensus. *Dig Dis*. 2019;37(4):266–83.
18. Gordon JP, McEwan PC, Maguire A, Sugrue DM, Puelles J. Characterizing unmet medical need and the potential role of new biologic treatment options in patients with ulcerative colitis and Crohn's disease: a systematic review and clinician surveys. *Eur J Gastroenterol Hepatol*. 2015;27(7):804–12.
19. Redlich K, Schett G, Steiner G, et al. Rheumatoid arthritis therapy after tumor necrosis factor and interleukin-1 blockade. *Arthritis Rheum*. 2003;48(12):3308–19.
20. Schwartz DM, Kanno Y, Villarino A, et al. JAK inhibition as a therapeutic strategy for immune and inflammatory diseases [published correction appears in *Nat Rev Drug Discov*. 2017;17(1):78]. *Nat Rev Drug Discov*. 2017;16(12):843–62.
21. Dhillon S, Keam SJ. Filgotinib: first approval. *Drugs*. 2020;80(18):1987–97.
22. Namour F, Diderichsen PM, Cox E, et al. Pharmacokinetics and pharmacokinetic/pharmacodynamic modeling of filgotinib (GLPG0634), a selective JAK1 inhibitor, in support of phase IIB dose selection. *Clin Pharmacokinet*. 2015;54(8):859–74.
23. Namour F. Author's reply to Srinivas: "Pharmacokinetics and pharmacokinetic/pharmacodynamic modeling of filgotinib (GLPG0634), a selective JAK1 inhibitor, in support of phase IIB dose selection." *Clin Pharmacokinet*. 2015;54(12):1297–8.
24. European Medicines Agency. EMA Jyseleca summary of product characteristics. https://www.ema.europa.eu/en/documents/product-information/jyseleca-epar-product-information_en.pdf. 2020. Accessed 7 Jun 2021.
25. Feagan BG, Danese S, Loftus EV Jr, et al. Filgotinib as induction and maintenance therapy for ulcerative colitis (SELECTION): a phase 2b/3 double-blind, randomised, placebo-controlled trial. *Lancet*. 2021;397(10292):2372–84.
26. European Medicines Agency. EMA Jyseleca Public Assessment Report. https://www.ema.europa.eu/en/documents/assessment-report/jyseleca-epar-public-assessment-report_en.pdf. 2020. Accessed 8 Nov 2021.

27. Genovese MC, Kalunian K, Gottenberg JE, et al. Effect of filgotinib vs placebo on clinical response in patients with moderate to severe rheumatoid arthritis refractory to disease-modifying antirheumatic drug therapy: the FINCH 2 randomized clinical trial. *JAMA*. 2019;322(4):315–25.
28. Food and Drug Administration. Testicular toxicity: evaluation during drug development guidance for industry. <https://www.fda.gov/media/117948/download>. 2018. Accessed 8 Jun 2021.
29. Amann RP. The cycle of the seminiferous epithelium in humans: a need to revisit? *J Androl*. 2008;29(5):469–87.
30. Cooper TG, Noonan E, von Eckardstein S, et al. World Health Organization reference values for human semen characteristics. *Hum Reprod Update*. 2010;16:231–45.
31. World Health Organization. WHO laboratory manual for the examination and processing of human semen (5th edn). <https://apps.who.int/iris/handle/10665/44261> 2010. Accessed Jan 2021.
32. Sikka SC, Chen C, Almas M, et al. Pregabalin does not affect sperm production in healthy volunteers: a randomized, double-blind, placebo-controlled, noninferiority study. *Pain Pract*. 2015;15(2):150–8.
33. Sikka SC, Hellstrom WJG. Current updates on laboratory techniques for the diagnosis of male reproductive failure. *Asian J Androl*. 2016;18(3):392–401.
34. Mountfield R, Bampton P, Prosser R, Muller K, Andrews JM. Fear and fertility in inflammatory bowel disease: a mismatch of perception and reality affects family planning decisions. *Inflamm Bowel Dis*. 2009;15(5):720–5.
35. Vermeire S, Schreiber S, Petryka R, et al. Clinical remission in patients with moderate-to-severe Crohn's disease treated with filgotinib (the FITZROY study): results from a phase 2, double-blind, randomised, placebo-controlled trial. *Lancet*. 2017;389(10066):266–75.
36. World Health Organization. WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction (4th edn). <https://www.aab.org/images/WHO%204th%20manual.pdf>. 1999. Accessed Jan 2021.
37. Carlsen E, Andersson AM, Petersen JH, Skakkebaek NE. History of febrile illness and variation in semen quality. *Hum Reprod*. 2003;18(10):2089–92.
38. Hellstrom WJG, Gittelman M, Jarow J, et al. An evaluation of semen characteristics in men 45 years of age or older after daily dosing with tadalafil 20mg: results of a multicenter, randomized, double-blind, placebo-controlled, 9-month study. *Eur Urol*. 2008;53(5):1058–65.
39. Jarow JP, Fang X, Hammad TA. Variability of semen parameters with time in placebo treated men. *J Urol*. 2013;189(5):1825–9.
40. European Medicines Agency. EMA Jyseleca European public assessment report. https://www.ema.europa.eu/en/documents/rmp-summary/jyseleca-epar-risk-management-plan-summary_en.pdf. 2020. Accessed 8 June 2021.
41. Leushuis E, van der Steeg JW, Steures P, et al. Semen analysis and prediction of natural conception. *Hum Reprod*. 2014;29(7):1360–7.
42. Wang C, Swerdloff RS. Limitations of semen analysis as a test of male fertility and anticipated needs from newer tests. *Fertil Steril*. 2014;102(6):1502–7.
43. Azenabor A, Ekun AO, Akinloye O. Impact of inflammation on male reproductive tract. *J Reprod Infertil*. 2015;16(3):123–9.
44. Rossato M, Foresta C. Antisperm antibodies in inflammatory bowel disease. *Arch Intern Med*. 2004;164(20):2283.
45. Shin T, Okada H. Infertility in men with inflammatory bowel disease. *World J Gastrointest Pharmacol Ther*. 2016;7(3):361–9.
46. World Health Organization. Sexual health and its linkages to reproductive health: an operational approach. <http://apps.who.int/iris/bitstream/handle/10665/258738/9789241512886-eng.pdf;jsessionid=BB192F998AC613EA23169F9D723943EE?sequence=1>. 2017. Accessed 3 Sept 2021.
47. Perez-Garcia LF, Micu M, Gheyle L, et al. Sexual function in male and female patients with rheumatoid arthritis: a post-hoc analysis of the FINCH studies. *Ann Rheum Dis*. 2021;80:496–7.