OPEN ACCESS

CLINICAL SCIENCE

Effects of filgotinib on semen parameters and sex hormones in male patients with inflammatory diseases: results from the phase 2, randomised, double-blind, placebo-controlled MANTA and MANTA-RAy studies

Walter Reinisch,[1] Wayne Hellstrom,[2] Radboud J E M Dolhain,[3] Suresh Sikka,[2] René Westhovens ,[4] Rajiv Mehta,[5] Timothy Ritter,[6] Ursula Seidler,[7] Oleksandr Golovchenko,[8] Vladimir Simanenkov,[9] Olena Garmish ,[10] Sławomir Jeka,[11] Radka Moravcová,[12] Vijay Rajendran,[13] Franck-Olivier Le Brun,[14] Sarah Arterburn,[15] Timothy R Watkins,[15] Robin Besuyen ,[16] Dirk Vanderschueren [17]

Handling editor Josef S Smolen

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/ard-2023-224017).

For numbered affiliations see end of article.

Correspondence to
Professor Dirk Vanderschueren, Department of Andrology and Endocrinology, KU Leuven, B-3001 Leuven, Belgium; dirk.vanderschueren@uzleuven.be

Received 13 February 2023
Accepted 19 April 2023

To cite: Reinisch W, Hellstrom W, Dolhain RJEM, *et al*. *Ann Rheum Dis* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/ard-2023-224017

ABSTRACT

Objectives The phase 2 MANTA and MANTA-RAy studies aimed to determine if the oral Janus kinase 1 preferential inhibitor filgotinib affects semen parameters and sex hormones in men with inflammatory diseases.

Methods MANTA (NCT03201445) and MANTA-RAy (NCT03926195) included men (21–65 years) with active inflammatory bowel disease (IBD) and rheumatic diseases (rheumatoid arthritis, spondyloarthritis or psoriatic arthritis), respectively. Eligible participants had semen parameters in the normal range per the WHO definition. In each study, participants were randomised 1:1 to receive once-daily, double-blind filgotinib 200 mg or placebo for 13 weeks for pooled analysis of the primary endpoint (proportion of participants with a ≥50% decrease from baseline in sperm concentration at week 13). Participants who met the primary endpoint were monitored over an additional 52 weeks for 'reversibility'. Secondary endpoints included change from baseline to week 13 in: sperm concentration, total motility, normal morphology, total count and ejaculate volume. Sex hormones (luteinising hormone, follicle stimulating hormone, inhibin B and total testosterone) and reversibility were exploratory endpoints.

Results Across both studies, 631 patients were screened, and 248 were randomised to filgotinib 200 mg or placebo. Baseline demographics and characteristics were similar within indications between treatment groups. Numerically similar proportions of filgotinib-treated versus placebo-treated patients met the primary endpoint (8/120 (6.7%) vs 10/120 (8.3%)), Δ−1.7% (95% CI −9.3% to 5.8%)). There were no clinically relevant changes from baseline to week 13 in semen parameters or sex hormones, or patterns of reversibility between treatment groups. Filgotinib was well tolerated, with no new safety events.

Conclusions Results suggest that once daily filgotinib 200 mg for 13 weeks has no measurable impact on semen parameters or sex hormones in men with active IBD or inflammatory rheumatic diseases.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Few data on the effects of advanced treatments for inflammatory bowel disease (IBD) and inflammatory rheumatic diseases (IRD) on male reproductive health are available.

WHAT THIS STUDY ADDS

⇒ The MANTA and MANTA-RAy studies were designed to determine the impact (if any) of filgotinib—a Janus kinase 1 preferential inhibitor—on semen parameters and sex hormones in men with active IBD or IRD.

⇒ Filgotinib 200 mg once daily for 13 weeks had no measurable impact on semen parameters or sex hormones.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This was a unique, landmark trial programme; the consistency of findings across its multiple endpoints demonstrates both the strength of the study design and the reliability of these results, hence it provides a methodological reference point for any future studies on this topic.

INTRODUCTION

Over the last two decades, major progress has been made in treating patients with inflammatory bowel disease (IBD) and inflammatory rheumatic diseases. Janus kinase (JAK) inhibitors are the most recent class of drugs approved for both indications. Filgotinib, an oral, JAK1 preferential inhibitor, is approved (at a dosage of 100 mg or 200 mg once daily) in the European Union (EU),[1] the UK,[2] Japan, South Korea and Taiwan as monotherapy or in combination with methotrexate (MTX), for the treatment of patients with moderate-to-severe rheumatoid arthritis (RA) who have an inadequate

response or intolerance to at least one disease-modifying antirheumatic drug. These approvals were based on results from the FINCH clinical trial programme.[3–5] Filgotinib (200 mg once daily) is also approved in the EU,[1] the UK[2] and Japan for the treatment of adults with moderately to severely active ulcerative colitis (UC), based on findings from the international phase 2b/3 SELECTION trial.[6] In addition, investigations into the efficacy and safety of filgotinib in patients with Crohn's disease (CD) are ongoing in the phase 3 DIVERSITY trial programme, and being initiated in patients with axial spondyloarthritis in the phase 3 OLINGUITO programme. Data from the exploratory, phase 2 DIVERGENCE2 trial suggest that future investigation in perianal fistulising CD may also be warranted.[7]

Few data on the effects of advanced treatments for IBD and inflammatory rheumatic diseases on male reproductive health are available.[8] Semen parameters, along with levels of serum gonadotropins (luteinising hormone (LH), follicle-stimulating hormone (FSH)) and testosterone, can be used to evaluate testicular function reliably. Other measures of reproductive health, such as pregnancy rates or testicular histology, are not feasible to assess in clinical trials.[9]

Based on preclinical findings in rats and dogs that demonstrated that male reproductive organs were affected by filgotinib (but not by its primary metabolite GS-829845) by an unknown mechanism,[10 11] the MANTA and MANTA-RAy studies were conducted. The two studies were conducted in consultation with global regulatory authorities to determine the impact (if any) of filgotinib 200 mg once daily on human semen parameters and sex hormones in participants with active IBD or inflammatory rheumatic diseases.

METHODS

Study design

MANTA and MANTA-RAy are multinational, randomised, double-blind (DB), placebo-controlled, parallel-group, phase 2 studies being conducted in patients with IBD (MANTA; ClinicalTrials.gov: NCT03201445; EudraCT: 2017-000402-38) and inflammatory rheumatic diseases (MANTA-RAy; ClinicalTrials.gov: NCT03926195; EudraCT: 2018-003933-14). MANTA was initiated in July 2017 in 11 countries; MANTA-RAy was initiated in May 2019 in 8 countries. Both studies were designed in line with guidance from the US Food and Drug Administration (FDA)[9] and in accordance with the Declaration of Helsinki, International Council for Harmonisation guidelines, and country-specific laws and regulations. Complete details of the design of both studies have been published previously.[10]

Each study included a 45-day screening period, during which each patient provided two semen samples for evaluation, the mean measurements of which were used to determine patient eligibility and to establish baseline semen parameter values. Eligible patients in each study were randomised 1:1 to receive filgotinib 200 mg or placebo daily for 13 weeks. Each study had its own randomisation list, and randomisation was stratified by the same stratification factors as used for the primary endpoint analyses detailed below.

In both studies, patients who experienced a ≥50% decrease compared with baseline in sperm concentration and/or percentage motility and/or percentage of sperm displaying normal morphology at week 13 (based on the mean of two evaluable samples), or at any subsequent visit while receiving treatment, prematurely discontinued study treatment regardless of whether the patient was a responder or non-responder to treatment in terms of their disease, and entered the monitoring phase for up to 52 weeks or until reversibility was observed, whichever occurred first. Reversibility was defined as all semen parameters that qualified the participant for entry into the monitoring phase returning to >50% of the baseline value. During the monitoring phase, patients received standard-of-care (SOC) therapy.

In MANTA, patients whose IBD responded to filgotinib 200 mg or placebo at week 13 continued to receive their assigned DB treatment through to week 26, while week 13 non-responders and patients with disease worsening between weeks 13 and 26 were switched to open-label (OL) filgotinib 200 mg.[10] Patients who were responders at week 26 or OL week 13 continued to receive DB treatment or OL filgotinib 200 mg, respectively, in the long-term extension (LTE) for up to 195 weeks. Non-responders at week 26/OL week 13 discontinued the study.

In MANTA-RAy, patients whose disease responded to filgotinib 200 mg or placebo at week 13 were unblinded and switched to OL filgotinib 200 mg or SOC therapy, respectively, in the extension phase up to week 156.[10] Non-responders were switched to SOC therapy. Choice of SOC therapy was at the investigator's discretion, as long as testicular toxicity was not included in the product label.

Patients in MANTA who experienced disease worsening during OL filgotinib treatment or during the LTE phase, and patients in MANTA-RAy who required treatment with a prohibited therapy, discontinued the study and attended an early termination visit. At the early termination visit, two semen samples were collected if semen samples had not been collected in the 2 weeks before early termination. This was followed by a safety visit 30 days after the last dose of study drug for patients who received filgotinib 200 mg or DB treatment.

Participants

Patients in both studies were men aged 21–65 years with an active inflammatory disease. In MANTA, patients had to have documented UC (minimum extent of 15 cm from the anal verge) or CD lasting at least 4 months, and had to meet the criteria for moderately to severely active IBD at screening. In MANTA-RAy, patients had to meet specific classification criteria for RA, ankylosing spondylitis, psoriatic arthritis or non-radiographic axial spondyloarthritis for at least 12 weeks before screening, and had to meet criteria for active disease during the screening period. Patients included in either study had to have the following measurements at screening based on the mean of two evaluable semen samples: sperm concentration ≥15 million/mL, sperm total motility ≥40%, normal sperm morphology ≥30%, total sperm per ejaculate ≥39 million and ejaculate volume ≥1.5 mL. Patients were excluded from either study if they had previously documented conditions affecting their reproductive health (eg, primary or secondary hypogonadism) or a previous diagnosis of reduced fertility. Full inclusion and exclusion criteria for each study and reasons for these, as well as details of permitted concomitant medications, are provided in the study design paper.[10] In brief, criteria were based on FDA guidance, which recommends values equal to or exceeding the generally accepted fifth percentile of WHO 2010 semen parameter reference values.[12]

Patient and public involvement

Neither study participants nor the public were involved in the design or conduct of MANTA or MANTA-RAy, or in the reporting or dissemination of the study data.

Endpoints and assessments

The primary endpoint for both studies was the proportion of participants with a ≥50% decrease from baseline in sperm concentration at week 13. This measure is a surrogate biomarker of testicular function, which is the FDA-recommended primary endpoint for studies of this type.[9] The study designs of MANTA and MANTA-RAy were similar up to week 13, which allowed for pooling of data until that time point; pooling was not possible thereafter. Week 13 was chosen for the primary endpoint because one spermatogenic cycle takes 13 weeks, and to limit the duration of treatment with placebo in MANTA-RAy to prevent irreversible, inflammatory disease-related damage in untreated patients with RA.

Secondary endpoints were the proportion of participants with a ≥50% decrease from baseline in sperm concentration at week 26, as well as the absolute and percentage changes from baseline in sperm concentration, sperm total motility, sperm morphology, total sperm count and ejaculate volume at week 13 and week 26.

Exploratory endpoints included monitoring of the reversibility of decreases of ≥50% from baseline in sperm concentration, motility and/or morphology up to the week 26/OL week 13 time point during the 52-week monitoring phase. Other exploratory endpoints were changes from baseline in sex hormones (LH, FSH, inhibin B and total testosterone) at week 13 and week 26. Further details on the rationale for the choice of primary, secondary and exploratory endpoints and the time points of these are provided in the study design paper.[10]

In both studies, semen samples were collected at screening and then every 13 weeks thereafter, as well as at early termination visits. For every semen assessment, two separate semen samples were collected within 14 days of each other. Samples had to be collected following an ejaculation-free period of at least 48 hours but no more than 7 days. Centralised morphology readings and stringent standardisation processes (with central oversight by the board-certified Andrology Director, Suresh Sikka and the staff at Tulane University, New Orleans, Louisiana, USA) were used to minimise intrareader and inter-reader variability and to ensure consistency across study sites.[10]

Safety evaluations (adverse events (AEs), haematology, serum chemistry, vital signs, body weight and symptom-directed physical examinations) and assessments of concomitant medication use were carried out at every study visit. Electrocardiograms and other laboratory measurements (urinalysis, lipids) were performed at predefined time points throughout each study. Complete details of procedure timings have been previously published.[10]

Statistical analyses

Based on FDA guidelines for studies on testicular toxicity,[9] it was agreed that a pooled sample size of 200 patients (100 patients per treatment group) across both studies would be adequate for providing cumulative distribution curves and an estimate with a reasonably narrow 95% CI for the primary endpoint. Because there is no well-established difference between treatment arms that would be considered clinically relevant in terms of the primary endpoint, MANTA and MANTA-RAy were not powered to detect predefined differences.[10]

Week 13 primary, secondary and exploratory endpoints (sex hormones) were analysed using the semen analysis set. For both studies, this set included all randomised patients who received at least one dose of DB study drug and who had two evaluable semen samples that were eligible for calculation of mean semen parameters at baseline and at the week 13 visit. Week 26 secondary endpoints were analysed using the MANTA week 26/LTE semen analysis set and the MANTA-RAy week 26 semen analysis set. The MANTA week 26/LTE semen analysis set included all patients who received at least one dose of study drug and who had two evaluable semen samples at baseline and at week 26/OL week 13. The MANTA-RAy week 26 semen analysis set included all patients who received at least one dose of OL filgotinib or SOC therapy in the extension phase and who had two evaluable semen samples at baseline and at week 26.

Cumulative safety data through to the data cut-off date (11 June 2021 for MANTA; 30 September 2021 for MANTA-RAy) for the submission were summarised. AEs in MANTA were analysed using the as-treated safety analysis set. This set included all randomised patients who received at least one dose of DB study drug. AEs and 'at-risk' periods were assigned based on the study drug that the patient received. AEs in patients in the placebo/OL filgotinib group in MANTA were counted in the placebo group if the AE onset was while the patient was receiving placebo, and counted in the filgotinib group if the onset was while the patient was receiving OL filgotinib. AEs that occurred during the DB treatment period in MANTA-RAy were analysed using the safety analysis set. This set included all randomised patients who received at least one dose of DB study drug. AEs in patients treated with OL filgotinib or SOC therapy in the MANTA-RAy extension phase were analysed using the MANTA-RAy extension phase safety analysis set.

Differences (filgotinib vs placebo) in the proportion of patients with a ≥50% decrease from baseline in sperm concentration at week 13 were calculated using the stratified Mantel-Haenszel method. Stratification factors for the pooled primary endpoint were disease type (IBD, inflammatory rheumatic disease) as denoted by the study (MANTA, MANTA-RAy), concurrent use of MTX (yes, no) and baseline sperm concentration (15–50 million/mL, >50 million/mL).

Changes from baseline in sperm concentration, sperm total motility, sperm morphology, total sperm count and ejaculate volume at week 13 were analysed using quantile regression. Differences between treatment groups in median changes from baseline in semen parameters at week 13 were calculated using quantile regression, adjusting for the baseline (continuous) value of the specific semen parameter of interest, the study (MANTA, MANTA-RAy), concurrent use of MTX (yes, no) and baseline sperm concentration as a categorical variable (15–50 million/mL, >50 million/mL). For sperm concentration, the baseline categorical sperm concentration variable was not needed in the model as baseline sperm concentration was already included. Proportions of patients with a ≥50% decrease from baseline in sperm concentration, and changes from baseline in each of the five semen parameters at week 26/OL week 13, were summarised descriptively by study and treatment sequence, owing to differences in the design of the two studies after week 13. Exact 95% CIs for the proportions of patients who had a ≥50% decrease from baseline in sperm concentration at week 26/OL week 13 by treatment sequence were based on the Clopper–Pearson method. Owing to the known variability and wide ranges of these variables, medians were assumed to be a more appropriate estimate than means for these data, because the distribution of data could have been skewed by outliers. The ClopperPearson method is explained in online supplemental material.

The number of patients who met reversibility criteria for a prespecified sperm parameter decrease (≥50% decrease in sperm concentration and/or motility and/or morphology) at week 13 and at week 26/OL week 13 was summarised descriptively.

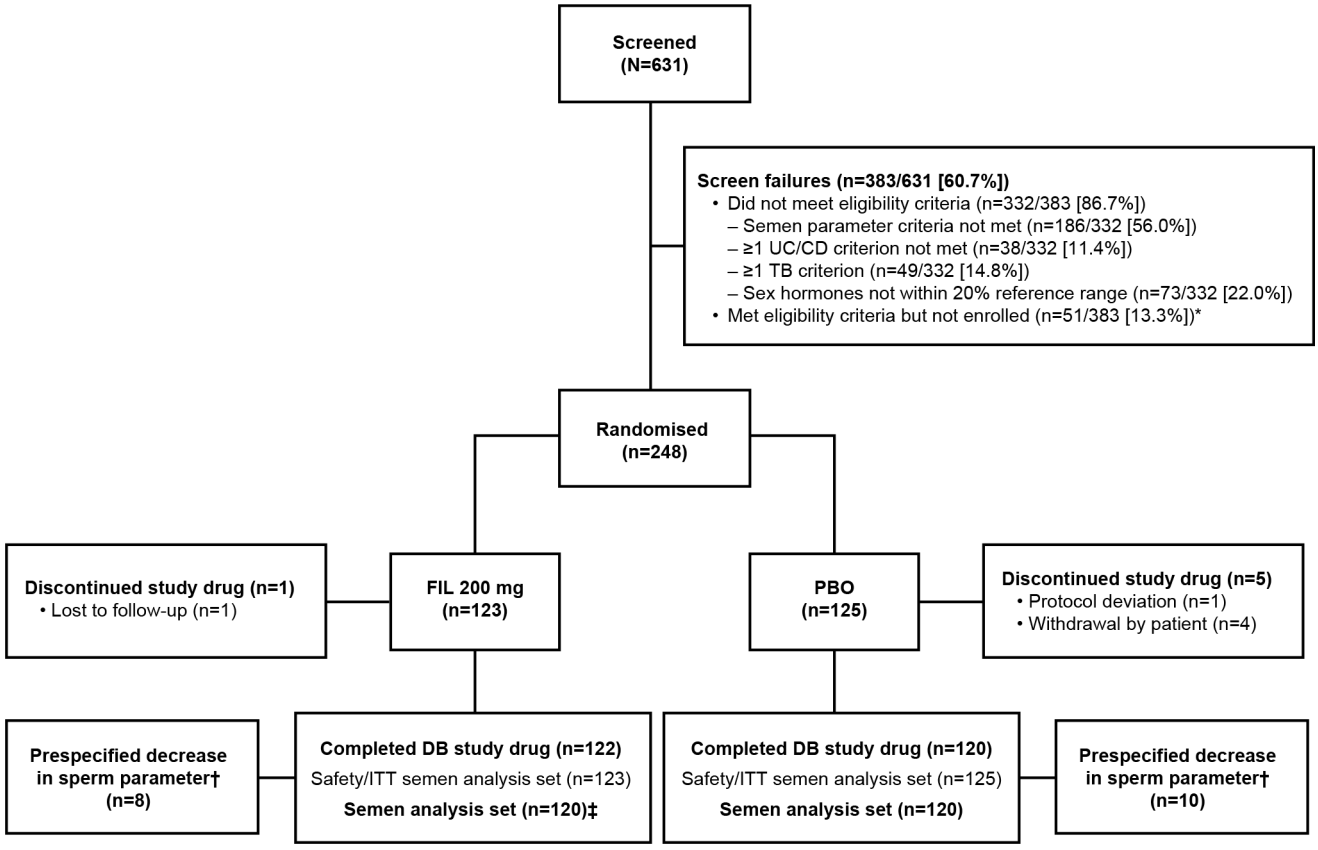

Figure 1 Patient disposition in MANTA and MANTA-RAy combined through to week 13. The reasons for screen failure were not mutually exclusive; only the four most frequently reported reasons for not meeting eligibility criteria are summarised. Patients who discontinued the study drug were those who discontinued before completing 13 weeks of DB treatment. *Reasons for patients not being enrolled, despite not violating at least one eligibility criterion: MANTA: semen sampling issues (7 patients), COVID-19-related national lockdown (5 patients), withdrew consent (4 patients), investigator's discretion (2 patients), sponsor's decision (2 patients), lost to follow-up (1 patient) and non-compliance (1 patient); MANTA-RAy: withdrew consent (12 patients), outside of visit window (6 patients), lost to follow-up (1 patient) and other reasons (10 patients). †Patients who had a prespecified decrease in any measured sperm parameters entered the monitoring phase. ‡Two patients did not have two evaluable semen samples at week 13 and were excluded from the semen analysis set. CD, Crohn's disease; DB, double-blind; FIL, filgotinib; ITT, intention-to-treat; PBO, placebo; TB, tuberculosis; UC, ulcerative colitis.

Changes from baseline in exploratory endpoints (sex hormones) at week 13 were analysed using quantile regression with a similar model to that used for change from baseline to week 13 in semen parameters (baseline hormone value was used in place of baseline semen parameter).

RESULTS

Patient flow and baseline characteristics

Across both studies, 631 patients were screened and 248 patients were randomised to receive filgotinib 200 mg or placebo (figure 1). In total, 332 patients (52.6%) did not meet the eligibility criteria, with the most common reason being that the semen parameter criteria were not met (n=186, 56.0%). Across the two studies, 240 patients (filgotinib, n=120; placebo, n=120) took at least one dose of study drug and had two evaluable semen samples at baseline and at week 13 and were included in the semen analysis set. Patient dispositions by individual treatment group through to week 26/OL week 13 are shown in online supplemental figure 1.

In the semen analysis set, the proportions of patients with UC, ankylosing spondylitis, psoriatic arthritis, non-radiographic axial spondyloarthritis, CD and RA were comparable across treatment groups (filgotinib: 50.8%, 25.8%, 11.7%, 4.2%, 4.2% and 3.3%, respectively; placebo: 52.5%, 26.7%, 12.5%, 0.8%, 3.3% and 4.2%, respectively).

Baseline demographics and disease characteristics were similar between treatment groups for patients included in the pooled week 13 semen analyses (table 1); in patients included in the semen analysis sets for the individual studies (online supplemental table 1); and for patients included in the week 26 semen analysis sets in each study (online supplemental table 2).

Primary endpoint

Pooled data

Numerically similar proportions of patients in MANTA and MANTA-RAy combined who were treated with filgotinib versus placebo had a ≥50% reduction from baseline in sperm concentration at week 13 (8/120 (6.7%) vs 10/120 (8.3%)), Δ −1.7% (95% CI −9.3% to 5.8%)). A cumulative distribution plot of percentage changes from baseline in sperm concentration at week 13 by DB treatment group is shown in figure 2. Approximately half of the patients in each treatment group had an increase, and the other half a decrease, in sperm concentration from baseline to week 13.

Although not prespecified as endpoints, cumulative distribution plots for sperm total motility and morphology showed

Table 1 Baseline demographics and disease characteristics among combined patients in MANTA and MANTA-RAy included in the week 13 analyses, by DB treatment group

	MANTA and MANTA-RAy	
	FIL 200 mg (n=120)	PBO (n=120)
Age* (years), mean±SD	38±8.7	37±8.5
Sex, male, n (%)	120 (100.0)	120 (100.0)
Race, n (%)		
American Indian or Alaska Native, Black or African American, or not permitted†	3 (2.5)	1 (0.8)
Asian	37 (30.8)	35 (29.2)
White	80 (66.7)	84 (70.0)
Ethnicity, n (%)		
Not Hispanic or Latino	118 (98.3)	118 (98.3)
Hispanic or Latino, or not permitted†	2 (1.7)	2 (1.7)
Type of inflammatory condition		
UC	61 (50.8)	63 (52.5)
AS	31 (25.8)	32 (26.7)
PsA	14 (11.7)	15 (12.5)
CD	5 (4.2)	4 (3.3)
RA	4 (3.3)	5 (4.2)
nrAxSpA	5 (4.2)	1 (0.8)
Duration of disease (years), mean±SD	5.4±5.08	5.5±5.17
MCS,‡ n (%)		
6	1/61 (1.6)	2/63 (3.2)
7	19/61 (31.1)	15/63 (23.8)
8	24/61 (39.3)	24/63 (38.1)
9	10/61 (16.4)	12/63 (19.0)
10	5/61 (8.2)	5/63 (7.9)
11	0/61	4/63 (6.3)
12	2/61 (3.3)	1/63 (1.6)
Mayo endoscopic subscore,‡ n (%)		
2	46/61 (75.4)	45/63 (71.4)
3	15/61 (24.6)	18/63 (28.6)
Crohn's DAI,§ mean±SD	290±29.3	281±5.7
PhGADA score,¶ mean±SD	72±14.1	70±14.1
Concomitant methotrexate use, n (%)	18 (15.0)	15 (12.5)
Smoking status, n (%)		
Former	8 (6.7)	18 (15.0)
Current	20 (16.7)	15 (12.5)
Never	92 (76.7)	87 (72.5)
Frequency of alcohol use in past 12 months, n (%)		
No alcohol in my life, or missing data	52 (43.3)	49 (40.8)
No alcohol in the last year, but drank in the past	10 (8.3)	11 (9.2)
1 or 2 times in the last year	8 (6.7)	4 (3.3)
3–11 times in the last year	14 (11.7)	9 (7.5)
1–3 times a month	28 (23.3)	23 (19.2)
Fewer than 4 times a week	8 (6.7)	18 (15.0)
4 or more times a week	0	4 (3.3)
Every day	0	2 (1.7)
Sperm concentration (M/mL), median (Q1, Q3)	61.7 (39.0, 85.5)	55.4 (36.4, 82.9)
Sperm concentration (M/mL), n (%)		
15–25	11 (9.2)	13 (10.8)
>25–50	37 (30.8)	40 (33.3)
>50	72 (60.0)	67 (55.8)
Total sperm count (M/ejaculate), median (Q1, Q3)	178.3 (111.3, 252.9)	166.8 (112.1, 259.0)
Ejaculate volume (mL), median (Q1, Q3)	2.9 (2.1, 3.9)	3.0 (2.0, 4.2)
Sperm total motility (%), median (Q1, Q3)	58.8 (50.0, 65.9)	58.3 (50.4, 64.4)
Sperm morphology (% normal), median (Q1, Q3)	41 (37, 45)	41 (37, 46)

Semen analysis set.
Baseline semen parameters were calculated based on the mean of two evaluable semen samples collected at the screening visit.
Sperm morphology was evaluated according to WHO 1992 criteria;[18] all other semen parameters were evaluated according to WHO 2010 criteria.[12] For other measurements, the baseline value was the last available value collected on or before the date of the first dose of DB study drug.
*Age was calculated from the date of the first dose of DB study drug.
†In the case that local regulators did not allow collection of race or ethnicity information.
‡Patients with UC only.
§Patients with CD only.
¶Patients with rheumatic diseases only.
AS, ankylosing spondylitis; CD, Crohn's disease; Crohn's DAI, Crohn's Disease Activity Index; DB, double-blind; FIL, filgotinib; M, millions; MCS, Mayo Clinic Score; nrAxSpA, non-radiographic axial spondyloarthritis; PBO, placebo; PhGADA, Physician's Global Assessment of Disease Activity; PsA, psoriatic arthritis; RA, rheumatoid arthritis; UC, ulcerative colitis.

a similar pattern (online supplemental figures 2 and 3). No filgotinib-treated patient had a ≥50% decrease from baseline in sperm motility or sperm morphology at week 13, and 1 placebo-treated patient had a ≥50% decline in all three sperm parameters at week 13.

Subgroup analyses of pooled data

Looking at stratification factors, numerically similar proportions of patients treated with filgotinib and placebo had a ≥50% reduction from baseline in sperm concentration at week 13 when stratified by study (MANTA, 1/66 (1.5%) vs 6/67 (9.0%), Δ −7.8% (95% CI −16.3% to 0.7%); MANTA-RAy, 7/54 (13.0%) vs 4/53 (7.5%), Δ 5.8% (95% CI −7.5% to 19.2%)).

Likewise, numerically similar proportions of patients treated with filgotinib and placebo met the primary endpoint when stratified by baseline sperm concentration, although decreases appeared more common among patients with higher baseline measures (15–50 million/mL, 1/48 (2.1%) vs 3/53 (5.7%); >50 million/mL, 7/72 (9.7%) vs 7/67 (10.4%)). There were no between-treatment group differences when stratified by disease (IBD, inflammatory rheumatic disease) or when stratified by concurrent use of MTX (yes, no).

Secondary endpoints

Pooled week 13 secondary endpoint data (median absolute and percentage changes from baseline in sperm concentration, sperm total motility, percentage normal sperm morphology, total sperm count and ejaculate volume) from MANTA and MANTA-RAy are shown in table 2. Median absolute and percentage changes from baseline in semen parameters at week 13 were mostly small and comparable between treatment groups. Similar results were observed in each of the individual studies (online supplemental table 3).

The proportions of patients with a ≥50% reduction from baseline in sperm concentration at week 26/OL week 13 are shown in online supplemental table 4. Proportions were low and showed no numerical differences across treatment groups. Median absolute and percentage changes from baseline in semen parameters measured at week 26/OL week 13 were small and

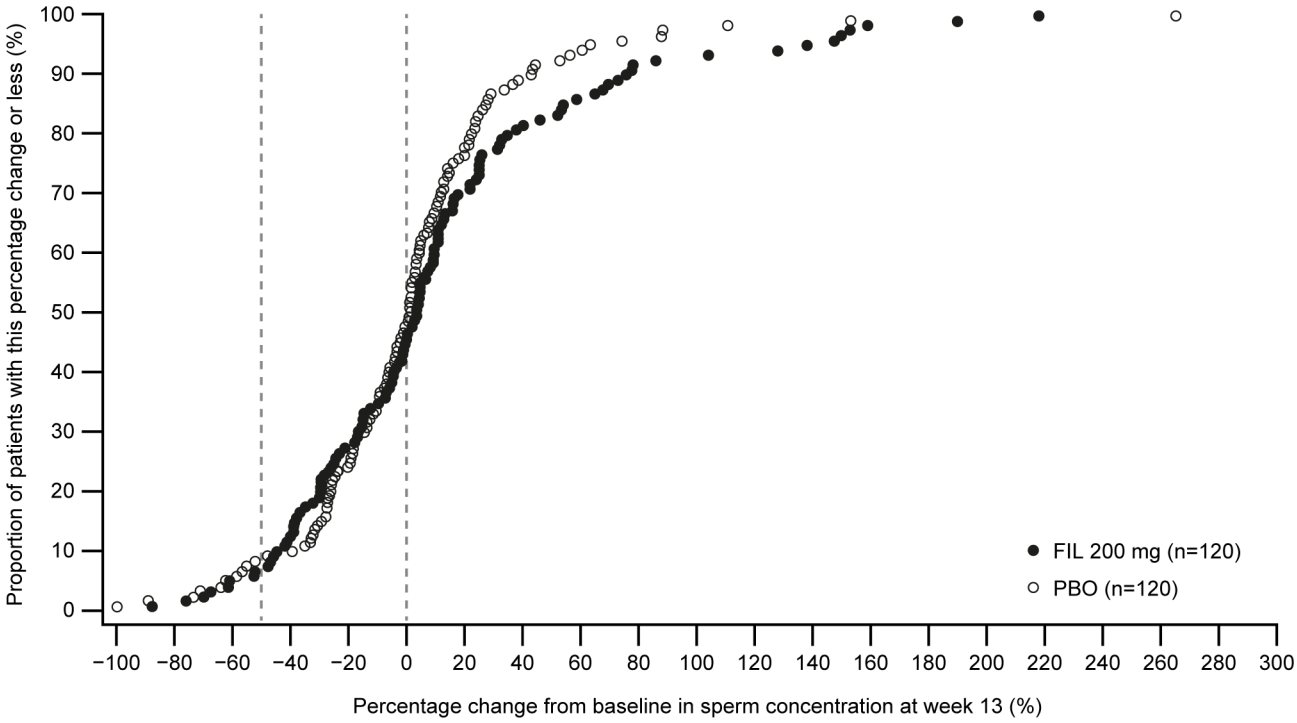

Figure 2 Percentage change from baseline in sperm concentration at week 13 among patients in MANTA and MANTA-RAy treated with FIL 200 mg or PBO. Semen analysis set. Baseline sperm/semen parameters were calculated based on the mean of two evaluable semen samples collected at the screening visit. Percentage change=((mean value at visit−mean value at baseline)/mean value at baseline)×100; the value at the visit was the mean of two evaluable samples collected at the visit. FIL, filgotinib; M, millions; PBO, placebo.

comparable between treatment groups and studies (online supplemental table 5).

Exploratory endpoints

Individual patient data on reversibility are shown in online supplemental table 6. Sixteen of 18 patients (6 of 8 filgotinib-treated and 10 of 10 placebo-treated) who entered the 52-week monitoring phase at week 13 demonstrated reversibility of prespecified sperm parameter reductions. One patient treated with filgotinib did not reverse by monitoring phase week 52. Reversibility for one patient treated with filgotinib was unknown, as the patient discontinued the study after monitoring phase week 13 before demonstrating reversibility. Eleven of 14 patients who entered the monitoring phase at week 26/OL week 13 demonstrated reversibility. Two filgotinib-treated patients (one filgotinib/filgotinib, one filgotinib/SOC) did not reverse at monitoring phase week 52. Reversibility for one additional patient treated with placebo was unknown, as he only provided one of the two required week 26 samples before withdrawing consent.

Changes from baseline in levels of sex hormones were small and not specific to any one hormone (online supplemental table 7). Patients with a ≥50% decrease in sperm concentration had no obvious corresponding changes in sex hormone levels.

AEs and changes in laboratory parameters and vital signs

Cumulative safety data were examined through the cut-off date. In MANTA, the mean (SD) exposure to DB or OL filgotinib among the 69 patients who were originally assigned to DB filgotinib was 73.1 (35.3) weeks. The mean (SD) exposure to DB placebo was 48.9 (38.2) weeks for the 70 patients who were originally assigned to DB placebo. Of these 70 patients, 23 switched to OL filgotinib, which they received for a mean (SD) duration of 55.4 (40.4) weeks. In MANTA-RAy, the mean (SD) exposure to the study drug among patients treated in the DB treatment phase was 16.6 (1.1) weeks in the filgotinib group and 16.1 (2.3) weeks in the placebo group. In the extension phase, among the patients who were initially treated with filgotinib, 39 patients continued to receive OL filgotinib for a mean (SD) duration of 59.7 (26.3) weeks, and 7 patients switched to SOC therapy for a mean (SD) duration of 73.3 (29.4) weeks. Among the 55 patients who were treated with placebo in the DB treatment phase, 49 patients went on to receive SOC treatment for a mean (SD) duration of 67.4 (23.6) weeks.

In both studies, the proportion of patients who experienced treatment-emergent AEs (TEAEs) was similar across the treatment groups (table 3). The most commonly reported TEAEs in each study are shown in online supplemental table 8. Serious AEs occurred in five patients in MANTA and in seven patients in MANTA-RAy (detailed in online supplemental material).

In terms of AEs of special interest, in MANTA, two patients had a serious infection while receiving DB study drug: one placebo-treated patient had a grade 3 AE of COVID-19 and one filgotinib-treated patient had a grade 3 AE of pneumonia. Two patients in MANTA-RAy experienced a serious infection (grade 3 AE of COVID-19 in filgotinib/OL filgotinib group; grade 3 AE of appendicitis in filgotinib/SOC group) during the extension phase. No events of herpes zoster or opportunistic infections (including active tuberculosis) were reported in any phase of either study. One patient in MANTA had a non-serious, grade 1 positive hepatitis B DNA assay result after switching from placebo to OL filgotinib, which led to discontinuation from the study per protocol. One filgotinib-treated patient in MANTA-RAy had a positive hepatitis B DNA assay result during the DB treatment phase. The event was non-serious and of grade 1 severity. The patient completed DB treatment but then discontinued the study per protocol owing to this AE. No malignancies,

Table 2 Change from baseline in semen parameters at week 13 among combined patients in MANTA and MANTA-RAy, by DB treatment group

	FIL 200 mg (n=120)	PBO (n=120)
Change from baseline in sperm concentration (M/mL)		
Baseline		
Median (Q1, Q3)	61.7 (39.0, 85.5)	55.4 (36.4, 82.9)
Week 13		
Median (Q1, Q3)	58.8 (38.8, 94.7)	51.9 (36.5, 83.8)
Change from baseline at week 13		
Median change (95% CI)	1.7 (−1.3, 4.7)	0.6 (−2.7, 2.1)
Difference in medians (FIL−PBO), M/mL (95% CI)*	3.0 (−1.3, 7.4)	–
% change from baseline at week 13		
Median % change (95% CI)	3.9 (−2.0, 9.6)	1.3 (−3.9, 4.2)
Difference in medians (FIL−PBO), % (95% CI)*	5.6 (−2.2, 13.3)	–
Change from baseline in % sperm total motility		
Baseline		
Median (Q1, Q3)	58.8 (50.0, 65.9)	58.4 (50.9, 64.5)
Week 13		
Median (Q1, Q3)	57.4 (49.4, 65.0)	56.9 (48.2, 68.6)
Change from baseline at week 13		
Median change (95% CI)	−0.7 (−1.8, 1.3)	−0.5 (−1.7, 1.1)
Difference in medians (FIL−PBO), % (95% CI)	−0.8 (−2.7, 1.2)	–
% change from baseline at week 13		
Median % change (95% CI)	−1.0 (−2.7, 1.9)	−0.9 (−2.6, 1.7)
Difference in medians (FIL−PBO), % (95% CI)	−1.3 (−4.7, 2.2)	–
Change from baseline in % normal sperm morphology		
Baseline		
Median (Q1, Q3)	41 (37, 45)	41 (37, 46)
Week 13		
Median (Q1, Q3)	43 (39, 46)	42 (39, 46)
Change from baseline at week 13		
Median change (95% CI)	1 (−1, 3)	1 (−1, 2)
Difference in medians (FIL−PBO), % (95% CI)	0 (−1, 1)	–
% change from baseline at week 13		
Median % change (95% CI)	2.7 (−1.1, 6.0)	2.3 (−1.0, 4.0)
Difference in medians (FIL−PBO), % (95% CI)	−0.3 (−2.8, 2.1)	–
Change from baseline in total sperm count (M/ejaculate)		
Baseline		
Median (Q1, Q3)	178.3 (111.3, 252.9)	166.8 (112.1, 259.0)
Week 13		
Median (Q1, Q3)	154.9 (93.9, 249.9)	153.9 (106.6, 216.4)
Change from baseline at week 13		
Median change (95% CI)	−10.0 (−19.8, 11.1)	−9.5 (−22.2, 2.4)
Difference in medians (FIL−PBO), M/ejaculate (95% CI)	3.8 (−13.6, 21.2)	–
% change from baseline at week 13		
Median % change (95% CI)	−4.7 (−11.3, 8.0)	−6.4 (−14.2, 1.8)
Difference in medians (FIL−PBO), % (95% CI)	0.2 (−11.0, 11.3)	–
Change from baseline in ejaculate volume (mL)		
Baseline		
Median (Q1, Q3)	2.9 (2.1, 3.9)	3.0 (2.0, 4.2)
Week 13		
Median (Q1, Q3)	2.9 (2.0, 3.6)	2.9 (2.0, 3.8)
Change from baseline at week 13		
Median change (95% CI)	−0.2 (−0.3, 0.1)	−0.1 (−0.3, 0.0)
Difference in medians (FIL−PBO), mL (95% CI)	0.0 (−0.2, 0.2)	–
% change from baseline at week 13		
Median % change (95% CI)	−6.9 (−10.3, 1.3)	−3.5 (−9.6, 0.0)
Difference in medians (FIL−PBO), % (95% CI)	−0.7 (−7.0, 5.5)	–

Semen analysis set.
Baseline sperm/semen parameters were calculated based on the mean of two evaluable semen samples collected at the screening visit.
Percentage change=((mean value at visit−mean value at baseline)/mean value at baseline)×100; the value at the visit was the mean of two evaluable samples collected at the visit.
Distribution-free 95% CI on median change and percentage change within group at week 13.
Difference in medians (FIL 200 mg—PBO), 95% CI on change and % change from baseline at week 13 is based on quantile regression adjusting for baseline value, study (for disease type (IBD, rheumatic disease)), concurrent methotrexate use (yes, no) and sperm concentration strata (15–50, >50 M/mL).
*Model for sperm concentration did not include sperm concentration strata (15–50, >50 M/mL).
DB, double-blind; FIL, filgotinib; IBD, inflammatory bowel disease; M, millions; PBO, placebo.

non-melanoma skin cancers, thromboembolic events, major adverse cardiac events or deaths were reported in any phase of either study.

Across both studies, changes in laboratory parameters were generally comparable across treatment groups and were not clinically relevant, with most being grade 1 or 2. Grades 3 and 4 changes in laboratory parameters are shown in online supplemental table 9. In addition, there were no clinically relevant changes from baseline in vital signs, body weight or body mass index, and no patients developed new clinically significant electrocardiogram abnormalities.

There were three pregnancies in the partners of patients in MANTA-RAy, all of which occurred in partners of patients in the filgotinib/OL filgotinib group, and three healthy babies were born. There were no reported pregnancies in the partners of patients in MANTA.

DISCUSSION

This is the first large-scale, placebo-controlled evaluation of the potential impact of an advanced therapy on semen parameters or sex hormones in patients with immune-mediated inflammatory diseases. The FDA-recommended primary endpoint was assessed based on pooled data from over 200 participants in two companion studies (MANTA and MANTA-RAy) with identical, randomised, placebo-controlled designs through to week 13. In line with FDA guidance, additional secondary and exploratory endpoints were evaluated, with the overall outcome of these studies determined using the totality of the week 13 and week 26 data. Overall, the observations reported in both studies provide no measurable evidence of an effect of filgotinib on testicular function.

There was no difference in the proportion of patients who had a ≥50% decrease from baseline in sperm concentration, or change from baseline in sperm concentration, sperm total motility, sperm morphology, total sperm count or ejaculate volume, between filgotinib and placebo-treated patients at week 13 in the pooled analysis. Furthermore, there were no apparent differences between treatment sequences at MANTA week 26/OL week 13 and MANTA-RAy week 26 in any semen parameter, or meaningful changes within treatment groups over time. Sex hormone (LH, FSH, inhibin B and total testosterone) levels were also generally similar between treatment groups, with any changes being intermittent and non-specific to any one hormone.

Few patients overall (8 filgotinib-treated and 10 placebo-treated) met the primary endpoint of a prespecified sperm

Table 3 Proportions of patients who experienced TEAEs in MANTA and MANTA-RAy, by treatment group

	MANTA		MANTA-RAy				
	As-treated		DB treatment phase		EXT phase		
No (%) of patients who experienced a TEAE	FIL 200 mg (n=92)	PBO (n=70)	FIL 200 mg (n=54)	PBO (n=55)	FIL/OL FIL responder (n=39)	FIL/SOC non-responder (n=7)	PBO/SOC responders and non-responders (n=49)
TEAEs	55 (59.8)	42 (60.0)	21 (38.9)	22 (40.0)	23 (59.0)	5 (71.4)	26 (53.1)
TEAEs of grade 3 or higher	5 (5.4)	2 (2.9)	2 (3.7)	1 (1.8)	2 (5.1)	1 (14.3)	2 (4.1)
Treatment-related TEAEs	4 (4.3)	6 (8.6)	9 (16.7)	6 (10.9)	10 (25.6)	0	3 (6.1)
Treatment-related TEAEs of grade 3 or higher	1 (1.1)	0	1 (1.9)	0	0	0	0
TE SAEs	3 (3.3)	2 (2.9)	2 (3.7)	1 (1.8)	3 (7.7)	1 (14.3)	1 (2.0)
Treatment-related TE SAEs	1 (1.1)	0	0	0	0	0	0
TEAEs leading to discontinuation of DB study drug	7/69 (10.1)	8/70 (11.4)	0	0	NA	NA	NA
TEAEs leading to discontinuation of OL FIL	4/47* (8.5)	0	NA	NA	5 (12.8)	NA	NA
TEAEs leading to discontinuation of study	11 (12.0)	8 (11.4)	1 (1.9)	0	2 (5.1)	0	0
Death	0	0	0	0	0	0	0

MANTA as-treated safety analysis set; MANTA-RAy safety/ITT semen analysis set and EXT phase safety analysis set.
AEs were coded using MedDRA V.24.0; severity grades were defined using CTCAE V.4.03.
*All four patients who discontinued OL FIL were originally randomised to PBO.
AE, adverse event; CTCAE, Common Terminology Criteria for Adverse Events; DB, double-blind; EXT, extension; FIL, filgotinib; ITT, intention-to-treat; MedDRA, Medical Dictionary for Regulatory Activities; NA, not applicable; OL, open-label; PBO, placebo; SAE, serious adverse event; SOC, standard of care; TE, treatment-emergent; TEAE, treatment-emergent adverse event.

concentration decrease at week 13. Of those, all patients demonstrated reversibility by the end of the monitoring phase, apart from one filgotinib-treated patient who had not reversed by week 52, and a second filgotinib-treated patient who withdrew consent prior to demonstrating reversibility. Similar results were seen among patients who entered the monitoring phase at week 26/OL week 13. A small imbalance towards filgotinib-exposed patients not demonstrating reversibility was observed; however, these data should be interpreted with caution, owing to the small numbers. Importantly, none of those patients demonstrated hormonal changes (in levels of testosterone/LH and FSH/inhibin B, which are reliable markers of testicular Leydig and Sertoli cell function, respectively)[13 14] that were suggestive of treatment-related primary testicular failure in any patient.

It is important to highlight that sperm concentration and other semen parameters are known to fluctuate over time.[15 16] Such known natural variation in semen quality is consistent with the balanced number of filgotinib-treated and placebo-treated patients with decreases of ≥50% in sperm concentration, the majority of which reversed over the course of the study. Natural variation may also explain the small number of patients demonstrating a lack of reversibility; these patients could have had regression to a normal mean from spuriously high baseline values. Indeed, decreases in sperm concentration seemed more frequent among patients with higher baseline values. Moreover, in each study, over half of the screened patients were not eligible for inclusion, with the main reason for screening failure being that the stringent semen parameter criteria were not met. Interestingly, similar screening failure rates and reasons have been observed in trials of healthy participants with similar entry criteria.[17] Together, the observed changes in semen parameters seem consistent with natural variation over time as opposed to an effect of disease or treatment; however, our findings cannot be extrapolated to patients with semen parameters beyond the applied eligibility criteria. Indeed, selection of patients with normal semen parameters at baseline adds to the robustness of the data, but limits the generalisability of these data to the overall population of patients with immune-mediated inflammatory diseases. These findings also cannot be extrapolated to directly make inferences about 'fertility', which encompasses a broader range of physiological functions than testicular function alone.

In terms of safety, the low discontinuation rate among patients treated with either filgotinib or placebo in both studies shows that filgotinib was well tolerated, and no new safety events were reported compared with other filgotinib trial programmes.

To our knowledge, very few studies on the effects of advanced treatments on semen parameters in patients with inflammatory diseases have been performed,[8] and to date there have been no other publications reporting the influence of JAK inhibitors on semen parameters in animals or humans. A major strength of these trials, which were designed according to FDA testicular toxicity investigation guidelines, is the parallel design through to week 13. This allowed pooling of the primary (and week 13 secondary) endpoint data to produce a large (240-patient) data set. One limitation is that the heterogeneity of study treatment arms after week 13 renders the interpretation of the data after this time point less robust. Importantly, patients provided two evaluable semen samples at each time point, to minimise intrapatient variability in semen parameters. Despite this, it is possible that there was non-adherence to the ejaculation-free period, which could have introduced bias. We aimed to reduce this in various ways, including reassessment if there was larger than expected variation in semen parameters between a patient's two samples. There are only a few biomarkers that can reliably monitor changes in human testicular function (such as semen analyses, serum testosterone and gonadotropin concentrations), and all of these were measured in this study. Notably, two additional semen parameters, progressive motility and vitality, were not measured here. This was because they are more complex to measure and hence less reliable than other semen parameters. Given the inherent variability in the semen parameters that were assessed, determining changes from baseline that are meaningful is difficult, and likely limited to extreme findings (another limitation of these studies). Nevertheless, a strength of these studies is the measurement of a combination of semen and hormone parameters, and the steps taken to ensure data were as robust as possible (details published previously).[10] It is of course

impossible to exclude all risk, but the totality of these data is not suggestive of a filgotinib-related effect on semen parameters.

CONCLUSIONS

The totality of the data from the MANTA and MANTA-RAy studies provides no evidence that results observed in rats and dogs were replicated in men, suggesting that filgotinib 200 mg once daily has no measurable impact on semen parameters, or on any other indicator of male reproductive health. The consistency of findings across endpoints in this unique, landmark trial programme demonstrates both the strength of the study design and the reliability of our results.

Author affiliations
[1]Medical University of Vienna, Department of Internal Medicine III, Division Gastroenterology & Hepatology, Vienna, Austria
[2]Tulane University Health Sciences Center, New Orleans, Louisiana, USA
[3]Department of Rheumatology, Erasmus University Medical Center Rotterdam, Rotterdam, Netherlands
[4]Skeletal Biology and Engineering Research Center, KU Leuven, Leuven, Belgium
[5]SIDS Hospital, Vijay Nagar, Gujarat, India
[6]GI Alliance, Southlake, Texas, USA
[7]Department of Gastroenterology, Hannover Medical School, Hannover, Germany
[8]Medical Clinical Investigational Center of Medical Center Health Clinic LLC, Vinnytsia, Ukraine
[9]State Budgetary Healthcare Institution "City Hospital #26", Saint-Petersburg, Russian Federation
[10]National Scientific Center M.D. Strazhesko Institute of Cardiology, Kyiv, Ukraine
[11]Clinic of Rheumatology and Systemic Connective Tissue Disorders, J. Biziel University Hospital No. 2, Bydgoszcz, Poland
[12]Department of Rheumatology, First Faculty of Medicine, Charles University and Rheumatology Institute, Prague, Czech Republic
[13]Galapagos NV, Mechelen, Belgium
[14]Galapagos GmbH, Basel, Switzerland
[15]Gilead Sciences, Inc, Foster City, California, USA
[16]Galapagos BV, Oegstgeest, Netherlands
[17]Laboratory of Clinical and Experimental Endocrinology, Department of Chronic Diseases, Metabolism and Ageing (CHROMETA), KU Leuven, Leuven, Belgium

Acknowledgements We would like to thank all investigators, site staff and patients for their contributions to this study (investigators are listed in the supplemental material). We would also like to thank the following study leads for their contributions: Goele Dekkers, William Barchuk, Dick de Vries, Lien Gheyle, Angi Gillen, Afsaneh Mozaffarian and Caroline Tonussi. Medical writing support for the preparation of this manuscript was provided by Frances Thompson, PhD, and Katie Pilldige, PhD, of PharmaGenesis London, London, UK, funded by Galapagos NV (Mechelen, Belgium). Publication coordination was provided by Slavka Baronikova, PhD, of Galapagos NV.

Contributors WH, SS, VR, TRW and RB contributed to study design. WR, WH, RJEMD, SS, RW, RMe, TR, US, OGo, VS, OGa, SJ, RMo, TRW and DV contributed to data collection. VR, FOLB, SA, TRW and RB contributed to data analysis. All authors contributed to data interpretation. All authors contributed to the development of the manuscript and approved the final version. All authors agree to be accountable for all aspects of the work. RB is the guarantor of the article.

Funding The MANTA study was sponsored by Gilead Sciences Inc. and conducted in collaboration with Galapagos NV. The MANTA-RAy study was sponsored by Galapagos NV and conducted in collaboration with Gilead Sciences Inc.

Competing interests WR has served as a speaker for AbbVie, Celltrion, Falk Pharma, Ferring, Janssen, Galapagos Medice, MSD, Roche, Pfizer, Pharmacosmos, Shire, Takeda, Therakos; as a consultant for AbbVie, Amgen, AOP Orphan, Arena Pharmaceuticals, Astellas, AstraZeneca, Bioclinica, Boehringer Ingelheim, Bristol Myers Squibb, Calyx, Celgene, Celltrion, Eli Lilly, Falk Pharma, Ferring, Galapagos, Gatehouse Bio, Genentech, Gilead, Grünenthal, ICON, Index Pharma, Inova, Janssen, Landos Biopharma, Medahead, MedImmune, Mitsubishi Tanabe Pharma Corporation, MSD, Novartis, OMass, Otsuka, Parexel, Periconsulting, Pharmacosmos, Pfizer, Protagonist, Provention, Quell Therapeutics, Sandoz, Seres Therapeutics, Setpointmedical, Sigmoid, Sublimity, Takeda, Teva Pharma, Therakos, Theravance, Zealand; as an advisory board member for AbbVie, Amgen, AstraZeneca, Boehringer Ingelheim, Bristol Myers Squibb, Celgene, Celltrion, Galapagos, Janssen, Mitsubishi Tanabe Pharma Corporation, MSD, Pharmacosmos, Pfizer, Sandoz, Takeda; and has received research funding from AbbVie, Janssen, MSD, Sandoz, Takeda. WH has been a consultant or adviser for Boston Scientific, Coloplast and Endo; has been an investigator for Coloplast and Endo; has been a lecturer for Endo; has been on advisory boards for Gilead Sciences, Maximus and Promescent; and is a board member, officer and trustee for Theralogix. RJEMD has received unrestricted research grants from the Dutch Arthritis Association, Galapagos and UCB; consulting fees from Galapagos; and speaking fees from AbbVie, Genzyme, Novartis, Lilly, Roche and UCB. SS declares no competing interests. RW has acted as a principal investigator, adviser and speaker for Celltrion, Galapagos and Gilead Sciences; and was an adviser for UCB. RMe declares no competing interests. TR has served on advisory boards or as a speaker for AbbVie, Arena Pharmaceuticals, Eli Lilly, Ferring Pharmaceuticals, Genentech, Gilead Sciences, Gossamer Bio, Intercept Pharmaceuticals, Janssen, Pfizer, Prometheus and Takeda. US has received grants from AbbVie, Abivax, Galapagos, Gilead Sciences, Index Pharmaceuticals, Janssen, Lilly, Roche-Genentech, Theravance Biopharma and Takeda; personal fees from AbbVie, Galapagos and Janssen; and non-financial support from Janssen and Galapagos. OGa declares no competing interests. VS declares no competing interests. OGo has served as a speaker for AbbVie, Amgen, Boehringer Ingelheim, Egis, Janssen, Johnson & Johnson, MSD, Novartis, Pfizer, Roche and Sandoz. SJ has received grants from Roche; consulting fees or honoraria from AbbVie,Amgen, Celgene, Lilly, Novartis, Roche, Sandoz, Sobi, UCB; and has been an advisory board member for Lilly, Novartis, Roche, Sandoz and Sobi. RMo declares no competing interests. VR, F-OLB and RB are employees and shareholders of Galapagos NV. TRW and SA are employees and shareholders of Gilead Sciences. DV received consultancy fees from Galapagos for the design and setup of the studies.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and these studies were conducted at multiple sites across 8 and 11 countries; trial protocols and amendments were approved by the relevant institutional review board at each participating site. Names and reference numbers are provided in online supplemental file uploaded. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Anonymised individual patient data will be shared on request for research purposes dependent upon the nature of the request, the merit of the proposed research, and the availability of the data and its intended use. The full data sharing policy for Gilead Sciences Inc., can be found at https://www.gilead.com/about/ethics-and-code-of-conduct/policies and for Galapagos NV can be found at https://www.clinicaltrials-glpg.com/us/en/data-transparency.html.

Author note A plain language summary is available for this article in the online supplemental material.

ORCID iDs
René Westhovens http://orcid.org/0000-0002-3432-3073
Olena Garmish http://orcid.org/0000-0001-8713-253X
Robin Besuyen http://orcid.org/0000-0002-8699-4549
Dirk Vanderschueren http://orcid.org/0000-0003-1395-0104

REFERENCES

1 European Medicines Agency. Filgotinib summary of product characteristics. Available: https://www.ema.europa.eu/en/medicines/human/EPAR/jyseleca#product-information-section [Accessed 09 Mar 2023].
2 Medicines and Healthcare products Regulatory Agency. Filgotinib summary of product characteristics. Available: https://www.medicines.org.uk/emc/product/11809/smpc [Accessed 25 Jan 2022].
3 Combe B, Kivitz A, Tanaka Y, *et al*. Filgotinib versus placebo or adalimumab in patients with rheumatoid arthritis and inadequate response to methotrexate: a phase III randomised clinical trial. *Ann Rheum Dis* 2021;80:848–58.
4 Genovese MC, Kalunian K, Gottenberg J-E, *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;322:315–25.
5 Westhovens R, Rigby WFC, van der Heijde D, *et al*. Filgotinib in combination with methotrexate or as monotherapy versus methotrexate monotherapy in patients with active rheumatoid arthritis and limited or no prior exposure to methotrexate: the phase 3, randomised controlled FINCH 3 trial. *Ann Rheum Dis* 2021;80:727–38.
6 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;397:2372–84.
7 Reinisch W, Colombel JF, D'Haens GR, *et al*. OP18 efficacy and safety of filgotinib for the treatment of perianal fistulizing Crohn's disease: results from the phase 2 divergence 2 study. J Crohns Colitis 2022;16:i019–21.
8 De Cock D, Brants L, Soenen I, *et al*. A systematic review on the effect of DMARDs on fertility in rheumatoid arthritis. *Semin Arthritis Rheum* 2020;50:873–8.
9 Food and Drug Administration. Testicular toxicity: evaluation during drug development guidance for industry. Available: https://www.fda.gov/media/117948/download [Accessed 08 Mar 2023].
10 Hellstrom WJG, Dolhain RJEM, Ritter TE, *et al*. MANTA and MANTA-RAy: rationale and design of trials evaluating effects of filgotinib on semen parameters in patients with inflammatory diseases. *Adv Ther* 2022;39:3403–22.
11 European Medicines Agency. EMA Jyseleca European public assessment report. Available: https://www.ema.europa.eu/en/documents/rmp-summary/jyseleca-epar-risk-management-plan-summary_en.pdf [Accessed 08 Mar 2023].
12 World Health Organization. *Laboratory manual for the examination and processing of human semen, 5th ed*. 2010.
13 O'Donnell L, Smith LB, Rebourcet D. Sertoli cells as key drivers of testis function. *Semin Cell Dev Biol* 2022;121:2–9.
14 Kristensen DM, Desdoits-Lethimonier C, Mackey AL, *et al*. Ibuprofen alters human testicular physiology to produce a state of compensated hypogonadism. *Proc Natl Acad Sci U S A* 2018;115:E715–24.
15 Mallidis C, Howard EJ, Baker HW. Variation of semen quality in normal men. *Int J Androl* 1991;14:99–107.
16 Keel BA. Within- and between-subject variation in semen parameters in infertile men and normal semen donors. *Fertil Steril* 2006;85:128–34.
17 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;53:1058–65.
18 World Health Organization. *Laboratory manual for the examination of human semen and sperm-cervical mucus interaction, 3rd ed*. 1992.