Non-Interventional Study Protocol

Study Protocol Number	MS700623_0049
Title	A non-interventional register-based comparative effectiveness study of rhFSH-alfa reference product vs. highly purified human menopausal gonadotropin or rhFSH-alfa biosimilar products for ovarian stimulation in *in vitro* fertilization or intracytoplasmic sperm injection treatment in Denmark and Sweden – The Nordic Follitropin Alfa Comparative Effectiveness (NORD-FACE) Study
Protocol version identifier	1.0
Protocol Date/Version	27 May 2021/Version 1.0
Replaces Version	10 May 2021/Version 1.0
EU PAS register number	Study not yet registered (to be registered in EU PAS register)
Active substance	G03GA05: Follitropin alfa G03GA02: Human menopausal gonadotropin
Medicinal product	Follitropin alfa, recombinant human follicle-stimulating hormone (rhFSH) Human menopausal gonadotropin
Sponsor	Merck Healthcare KGaA an affiliate of Merck KGaA, Darmstadt, Germany Frankfurter Str. 250 Darmstadt, Germany
Research question and objectives	The research question is whether, among women undergoing in vitro fertilization (IVF) or intracytoplasmic sperm injection (ICSI) treatment, recombinant human follicle-stimulating hormone (rhFSH)-alfa reference product is associated with better treatment results than highly purified human menopausal gonadotropin (HP-hMG) or rhFSH-alfa biosimilar products in terms of clinical effectiveness and safety outcomes. **Primary Objective:** The primary objective of this study is to compare rhFSH-alfa reference product with HP-hMG or rhFSH-alfa biosimilar products regarding the following live birth outcomes: a) Live birth rate per initiated IVF/ICSI stimulation cycle b) Cumulative live birth rate per initiated IVF/ICSI stimulation cycle c) Cumulative live birth rate in up to five initiated IVF/ICSI stimulation cycles **Secondary Objectives:** To compare rhFSH-alfa reference product with HP-hMG or rhFSH-alfa biosimilar products regarding the following outcomes: 1. a) Clinical pregnancy rate and ongoing pregnancy rate per initiated IVF/ICSI stimulation cycle, and b) Cumulative clinical pregnancy rate and cumulative ongoing pregnancy rate per initiated IVF/ICSI stimulation cycle

2. Number of oocytes retrieved, embryos transferred, embryos cryopreserved, and utilizable embryos per initiated IVF/ICSI stimulation cycle and per oocyte retrieval cycle
3. Implantation rate
4. Time-to-live birth in up to five IVF/ICSI stimulation cycles
5. Rate of pregnancy loss for the first (and successive) IVF/ICSI stimulation cycle(s)
6. Rate of multiple pregnancy for the first (and successive) IVF/ICSI stimulation cycle(s)
7. Number of cancelled cycles for the first (and successive) IVF/ICSI stimulation cycle(s)
8. Number of ovarian hyperstimulation syndrome cases for the first (and successive) IVF/ICSI stimulation cycle(s)

An additional secondary objective describes costs associated with the rhFSH-alfa reference product, HP-hMG and rhFSH-alfa biosimilar products individually and overall, across all products:

9. Costs associated with IVF/ICSI treatment (including drugs and other treatments), miscarriage/birth, and adverse events (i.e., OHSS)

Countries of study	Denmark Sweden
Author	PPD [redacted]

[redacted]

1 Table of Contents

2 List of Abbreviations

ART	Assisted Reproductive Technology
ATC	Anatomical Therapeutic Chemical Classification System
BMI	Body Mass Index
CCPR	Cumulative Clinical Pregnancy Rate
CPR	Clinical Pregnancy Rate
CLBR	Cumulative Live Birth Rate
COPR	Cumulative Ongoing Pregnancy Rate
COS	Controlled Ovarian Stimulation
COVID-19	Coronavirus Disease 2019
CT	Clinical Trail
DDD	Defined Daily Dose
DRG	Diagnosis Related Group
DIVF	Danish In-Vitro Fertilization Register
DMBR	Danish Medical Birth Register
DNPR	Danish National Patient Register
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
ESHRE	European Society of Human Reproduction and Embryology
EU	European Union
EU PAS register	European Union Electronic Register of Post-Authorization Studies
FET	Frozen Embryo Transfer
FSH	Follicle-Stimulating Hormone
GnRH	Gonadotropin-Releasing Hormone
GP	General Practitioner
GPP	Good Pharmacoepidemiology Practices
GVP	Good Pharmacovigilance Practices
hCG	Human Chorionic Gonadotropin
HP-hMG	Highly Purified Human Menopausal Gonadotropin

hMG	Human Menopausal Gonadotropin
ICD-9	International Classification of Diseases, 9th revision
ICD-10	International Classification of Diseases, 10th revision
ICER	Incremental Cost-Effectiveness Ratio
ICMART	International Committee for Monitoring Assisted Reproductive Technologies
ICMJE	International Committee of Medical Journal Editors
ICSI	Intracytoplasmic Sperm Injection
INN	International Nonproprietary Name
IPTW	Inverse Probability of Treatment Weighting
IU	International Unit
IUI	Intrauterine Insemination
IVF	In Vitro Fertilization
L	Liter
LBR	Live Birth Rate
LH	Luteinizing Hormone
LPS	Luteal Phase Support
MA	Meta-Analysis
MC	Multiple-Cycle
MC-CLBR	Multiple-Cycle Cumulative Live Birth Rate
mIU	milli-International Unit
NCMP	Nordic Classification of Medical Procedures
NCSP	Nordic Classification of Surgical Procedures
NIS	Non-Interventional Study
N/A	Not Applicable
NOMESCO	Nordic Medico-Statistical Committee
NPU	Nomenclature for Property and Unit
OHSS	Ovarian Hyperstimulation Syndrome
OPR	Ongoing Pregnancy Rate
OR	Odds Ratio
PCOS	Polycystic Ovarian Syndrome
pmol	Picomole

PS	Propensity Score
QMS	Quality Management System
Q-IVF	Swedish National Quality Registry of Assisted Reproduction
RCT	Randomized Clinical Trial
rhFSH	Recombinant Human Follicle-Stimulating Hormone
RLRR	Register of Laboratory Results for Research
RMPS	Danish Register of Medicinal Product Statistic
RWD	Real-World Data
RWE	Real-World Evidence
SAP	Statistical Analysis Plan
SID	Study Identification Number
SD	Standard Deviation
SMBR	Swedish Medical Birth Register
SNPR	Swedish National Patient Register
SPDR	Swedish Prescribed Drugs Register
TTLB	Time-to-Live Birth
TTP	Time-to-Pregnancy
USA	United States of America
WHO	World Health Organization

3 Responsible Parties

Responsible parties	Contact details
Sponsor	Merck Healthcare KGaA an affiliate of Merck KGaA, Darmstadt, Germany Frankfurter Str. 250 Darmstadt, Germany
Sponsor contact person (Protocol Lead)	PPD PPD EMD Serono: A business of Merck KGaA, Darmstadt, Germany EMD Serono Research & Development Institute, Inc. 45A Middlesex Turnpike, Billerica, MA 01821, USA
Coordinating center	PPD PPD PPD
Principal-in-charge	PPD PPD
Principal investigator	PPD PPD
Other investigators in the coordinating center	PPD PPD PPD PPD PPD PPD PPD

Responsible parties	Contact details
	PPD PPD
Key Opinion Leader / Clinical Expert: Denmark	PPD PPD
Key Opinion Leader / Clinical Expert: Sweden	PPD PPD
Investigators who have substantially contributed to the study conception or design	PPD

4 Abstract

Title	A non-interventional register-based comparative effectiveness study of rhFSH-alfa reference product vs. highly purified human menopausal gonadotropin or rhFSH-alfa biosimilar products for ovarian stimulation in in vitro fertilization or intracytoplasmic sperm injection treatment in Denmark and Sweden, 2010-2020 – The Nordic Follitropin Alfa Comparative Effectiveness (NORD-FACE) Study Version 1.0, Date: 27 May 2021 Main author: PPD
Rationale and background	Gonadotropins extracted from the urine of post-menopausal women were the first drugs used to stimulate folliculogenesis in the treatment of infertility and in assisted reproductive technology (ART). This old generation of human menopausal gonadotropin (hMG) consists not only of a mixture of follicle-stimulating hormone (FSH), luteinizing hormone (LH) and human chorionic gonadotropin (hCG), but also may include biologically active contaminants. The last generation of gonadotropins is represented by highly purified recombinant products; human FSH (rhFSH) products, including follitropin alfa. These rhFSH products are used for ovulation induction in anovulatory women and for stimulation of multifollicular development in women undergoing superovulation for ART, including intrauterine insemination. Follitropin alfa (GONAL-f®), in this study termed the "rhFSH-alfa reference product", was first approved in 1995. Currently, two rhFSH-alfa biosimilar products are available on the European Union (EU) market (Ovaleap®, authorized since 2013; Bemfola®, authorized since 2014). Approval of biosimilars has the main objective to guarantee treatment accessibility to all the patients, in an ideal case, due to lower price but with the same safety and efficacy as the reference product. The rhFSH-alfa reference product has been compared to both hMG and rhFSH-alfa biosimilars with respect to pregnancy outcomes. However, available evidence originates from a few randomized clinical trials (RCTs) powered for pregnancy outcomes and mainly evidence has been limited by the short duration of study follow-up and the fact that live birth and other meaningful endpoints, such as ongoing pregnancy rate and clinical pregnancy, have been considered as secondary outcomes in trials and, in turn, are usually only partially reported. Accordingly, previous trial evidence has yielded mixed results or lacked sufficient statistical power to enable meaningful comparisons among treatment groups for these secondary parameters. Therefore, the use of real-world data (RWD) is an efficient way to address these types of long-term outcomes, as patient data are collected longitudinally, over a longer period of time, and stem from routine clinical practice. Moreover, clinical evidence generated based on RWD and the associated cost of treatments, is valuable for

	health-economic evaluations and provides unique evidence for healthcare decision makers from different sectors, which can improve patient access and reduce healthcare burden. The hypothesis tested in this study is that there are possible differences in effectiveness and safety outcomes between rhFSH-alfa reference product and urinary gonadotropins or rhFSH-alfa biosimilar products, with a beneficial effect of rhFSH-alfa reference product in terms of meaningful clinical outcomes. With the long-term market use of rhFSH-alfa reference and urinary products, and the increased use of rhFSH-alfa biosimilar products, it is important to assess this hypothesis in the routine clinical practice setting. This will be achieved by comparing the effectiveness, and safety, of rhFSH-alfa reference product vs. highly purified (HP-)hMG or rhFSH-alfa biosimilar products, in different GnRH down-regulation regimens, regarding (cumulative) live birth and other effectiveness and safety outcomes that would be relevant for physicians, and most importantly for patients.

Research question and objectives	The research question is whether rhFSH-alfa reference product is associated with better treatment results than HP-hMG or rhFSH-alfa biosimilar products in terms clinical effectiveness and safety outcomes.
	
	Primary Objective:
	
	• The primary objective of this study is to compare rhFSH-alfa reference product with HP-hMG or rhFSH-alfa biosimilar products regarding the following live birth outcomes:
	  ◦ live birth rate (LBR) per initiated in vitro fertilization (IVF) or intracytoplasmic sperm injection (ICSI) stimulation cycle,
	  ◦ cumulative live birth rate (CLBR) per initiated IVF/ICSI stimulation cycle
	  ◦ CLBR in up to five initiated IVF/ICSI stimulation cycles [termed multiple-cycle (MC-)CLBR]
	
	Secondary Objectives:
	
	• There are multiple secondary objectives comparing rhFSH-alfa reference product with HP-hMG or rhFSH-alfa biosimilar products with respect to the following outcomes:
	
	1. a) Clinical pregnancy rate (CPR) and ongoing pregnancy rate (OPR) per initiated IVF/ICSI stimulation cycle, and b) cumulative clinical pregnancy rate (CCPR) and cumulative ongoing pregnancy rate (COPR) per initiated IVF/ICSI stimulation cycle

	2. Number of oocytes retrieved, embryos transferred, embryos cryopreserved, and utilizable embryos per initiated IVF/ICSI stimulation cycle and per oocyte retrieval cycle 3. Implantation rate 4. Time-to-live birth (TTLB) 5. Rate of pregnancy loss for the first (and successive) IVF/ICSI stimulation cycle(s) 6. Rate of multiple pregnancy for the first (and successive) IVF/ICSI stimulation cycle(s) 7. Number of cancelled cycles for the first (and successive) IVF/ICSI stimulation cycle(s) 8. Number of ovarian hyperstimulation syndrome (OHSS) cases for the first (and successive) IVF/ICSI stimulation cycle(s) An additional secondary objective describes treatment-associated costs. Hence, the secondary objective 9 is to describe the following outcomes for the rhFSH-alfa reference product, HP-hMG and rhFSH-alfa biosimilar products individually and overall, across all products: 9. Costs associated with IVF/ICSI treatment, miscarriage/birth, and adverse events (i.e., OHSS)
Study Design	This is a non-interventional study based on secondary data from national population-based registers with prospective data collection in Denmark and Sweden. The study uses a cohort design and is conducted as a comparative effectiveness and safety study with head-to-head comparisons of drugs used for treatment of infertility and in ART. The study drugs are rhFSH-alfa reference product (drug of interest), HP-hMG and rhFSH-alfa biosimilar products (comparator drugs). The overall study period will be between 2010 and 2020, and different study period starts will be applied depending on which two drugs are being compared.
Population	Inclusion criteria: • Initiated IVF/ICSI stimulation cycle with rhFSH-alfa reference product, HP-hMG or rhFSH-alfa biosimilar product monotherapy for controlled ovarian stimulation (COS) during the study period • Aged 18 years or more at stimulation cycle index date • Female sex at stimulation cycle index date

	Exclusion criteria: • History of 5 or more IVF/ICSI stimulation cycles prior to the stimulation cycle index date • Ovarian stimulation for the purpose of oocyte donation, oocyte storage, embryo donation or oncological or other medically indicated embryo storage, or preimplantation genetic testing • Non-availability of individual-level data on IVF/ICSI treatment, dispensed drugs and medical history for 36 months or more prior to and including the stimulation cycle index date
Outcomes	The primary outcome of this study is live birth, measured as LBR, CLBR, MC-CLBR (primary objective) and TTLB (secondary objective 4). There are several secondary outcomes to be assessed, which are connected to the secondary study objectives and include: clinical pregnancy, ongoing pregnancy, oocytes retrieved, embryos transferred, embryos cryopreserved, utilizable embryos, implantation, pregnancy loss, multiple pregnancy, cycle cancellation, OHSS, and treatment-associated costs.
Variables	Exposure variables: rhFSH-alfa reference product (follitropin alfa, ATC code G03GA05), HP-hMG (ATC code G03GA02) or rhFSH-alfa biosimilar product (ATC code G03GA05) used for COS will be established based on information primarily from prescriptions dispensed in community pharmacies. Other covariates include demographic characteristics, IVF/ICSI treatment characteristics, other clinical characteristics, fertility-related medical history, and laboratory test results (available in Denmark only).
Data Sources	Secondary high-quality data from population-based registers with national coverage in Denmark and Sweden will be used in the study. Danish data sources include the Civil Registration System (CPR), the Danish In-vitro Fertilization Register (DIVF), the Danish National Patient Register (DNPR), the Danish Medical Birth Register (DMBR), the Danish Register of Medicinal Product Statistics (RMPS), and The Danish Register of Laboratory Results for Research (RLRR). Swedish data sources include the Swedish National Quality Registry of Assisted Reproduction (Q-IVF), the Swedish Medical Birth Register (SMBR), the Swedish National Patient Register (SNPR), the Swedish Prescribed Drug Register (SPDR), and the Swedish Cause of Death Register. All study permit approvals and access to the study data will be applied for by PPD, and the application for access to individual-level data for scientific research follows a standard application procedure at the Danish and Swedish authorities. Before PPD can access the data, the data

	holders have collected and managed data according to their own standards. Once the data can be accessed, PPD will start processing the data and will maintain information on the study individuals securely on site according to up-to-date standard operating procedures.
Study Size	The study size was estimated for the primary outcome of live birth (measured as LBR, CLBR and MC-CLBR) based on publicly available annual reports from the Danish and Swedish IVF registers and the previous feasibility study conducted by PPD . As each woman to be included in the analysis will have at least one stimulation cycle, the number of women having a stimulation cycle with a study drug is the limiting factor for the study size. Thus, the study size was estimated based on the number of women and using the outcome measure of MC-CLBR. The anticipated approximate maximum number of women who have initiated a stimulation cycle with a study drug, and will be available for the comparison of rhFSH-alpha reference product and HP-hMG products (study period starting in 2010), is as follows: • 19,000 women in rhFSH-alpha reference product cohort and 12,000 women in HP-hMG cohort, in Denmark • 18,000 women in rhFSH-alpha reference product cohort and 17,000 women in HP-hMG cohort, in Sweden For the comparison of rhFSH-alfa reference product and rhFHS-alfa biosimilar products (study period starting in 2014), the corresponding numbers are the following: • 14,000 women in rhFSH-alpha reference product cohort and 3,000 women in rhFSH-alpha biosimilar product cohort, in Denmark • 11,000 women in rhFSH-alpha reference product cohort and 10,000 women in rhFSH-alpha biosimilar product cohort, in Sweden The conservative minimal detectable difference between the study groups was investigated to show the minimal differences in primary outcomes that can be detected in this study. The minimal detectable difference was calculated by using the measure of association in this study: odds ratio (OR) and its corresponding 95% confidence interval (CI). The minimal detectable difference is the difference in the outcome rates between the study cohorts that will yield an OR whose lower 95% CI is above 1. To take into account the effect the study inclusion and exclusion criteria, a crude assumption that 75% of the women are eligible for the study. With this assumption, the anticipated study size will be able to show any differences of 1.5 percentage point (% point) in primary outcomes between the rhFSH-alfa reference product and HP-hMG study cohorts, in both Denmark and Sweden. For comparison of rhFSH-alfa reference

	product and biosimilars, the anticipated minimum difference that can be detected in primary outcomes will be 2.5% point in Denmark and 2.0% point in Sweden.
Data Analysis	To address the study objectives, analysis sets will be constructed based on the associated analyses. For example, three analysis sets will be constructed for the primary objective: • Stimulation cycles with drug of interest initiated on or before 31 December 2019 (LBR) • Stimulation cycles with drug of interest and linked FET cycles initiated on or before 31 December 2019 (CLBR) • Women who initiated their first-ever stimulation cycle with drug of interest on or before 31 December 2019 (MC-CLBR) The main data analysis will be conducted in two stages: (i) construction of the inverse probability of treatment weighted (IPTW) study cohorts by modelling the rhFSH-alfa reference product vs. HP-hMG treatment initiation and rhFSH-alfa reference product vs. rhFSH-alfa biosimilar product treatment initiation, (ii) estimating the effectiveness of rhFSH-alfa reference product on the (cumulative) live birth rates using adjusted (IPTW-weighted) odds ratios (ORs), comparing rhFSH-alfa reference product study cohort to the comparators (HP-hMG study cohort or rhFSH-alfa biosimilars study cohort). Descriptive analysis will be conducted to describe the baseline characteristics of the study cohorts. For outcomes, rates per 100 units of observations will be estimated with 95% CIs. Adjusted ORs with 95% CIs will be estimated from the statistical model weighted with IPTWs and any variables included in the propensity score (PS) that are still unbalanced between the study cohorts after weighting. All analyses will be performed separately for each country. A meta-analysis approach based on aggregate data may be considered for the primary objective to provide summary estimates with increased precision based on the entire study population across both countries. Relevant sensitivity analyses will be performed to explore the robustness of the results from the main analysis of the primary objective.
Milestones	Registration in the EU PAS register — May 2021 Start of data permit process — March 2021 End of data permit process — December 2021 Start of data extraction (for Denmark) — August 2021 End of data extraction (for Denmark) — September 2021 Start of data extraction (for Sweden) — January 2022 End of data extraction (for Sweden) — February 2022

	Start of *interim* data analysis	September 2021
	End of *interim* data analysis	December 2021
	Start of *interim* study reporting process	January 2022
	Interim report of study results	April 2022
	Start of *final* data analysis	March 2022
	End of *final* data analysis	June 2022
	Start of *final* study reporting process	June 2022
	Final report of study results	July 2022
	Start of manuscript writing process	September 2022
	Final manuscript delivery	November 2022

5 Amendments and Updates

None.

6 Milestones

The following table estimates timelines for the milestones of this multi-country study, combining results from each country.

Milestone	Planned date
Registration in the EU PAS register	May 2021
Start of data permit process	March 2021
End of data permit process	December 2021
Start of data extraction (for Denmark)	August 2021
End of data extraction (for Denmark)	September 2021
Start of data extraction (for Sweden)	January 2022
End of data extraction (for Sweden)	February 2022
Start of *interim* data analysis[+]	September 2021
End of *interim* data analysis[+]	December 2021
Start of *interim* study reporting process[+]	January 2022
Interim report of study results[+]	April 2022
Start of *final* data analysis	March 2022
End of *final* data analysis	June 2022
Start of *final* study reporting process	June 2022
Final report of study results	July 2022
Start of manuscript writing process	September 2022
Final manuscript delivery	November 2022

[+]Includes data from Denmark only (initially expected to arrive prior to Swedish data). Interim analysis and interim report are currently tentative due to a recently extended data application processing time for Denmark (due to COVID-19).

7 Rationale and Background

Gonadotropins extracted from the urine of post-menopausal women were the first drugs used to stimulate folliculogenesis in the treatment of infertility and in assisted reproductive technology (ART). This old generation of human menopausal gonadotropin (hMG) consists not only of a mixture of follicle-stimulating hormone (FSH), luteinizing hormone (LH) and human chorionic gonadotropin (hCG), but also may include biologically active contaminants, such as growth factors, binding proteins and prion proteins (1).

The last generation of gonadotropins is represented by highly purified recombinant products. These recombinant human FSH (rhFSH) products, including follitropin alfa, are used for ovulation induction in anovulatory women and for stimulation of multifollicular development in women undergoing superovulation for ART, including intrauterine insemination (IUI). Follitropin alfa (GONAL-f®), hereafter referred to as "rhFSH-alfa reference product," was first approved in 1995 (2).

The urinary gonadotropins had been used universally until the introduction of recombinant technology. Even if this technology had shown improvement in purity, consistency and specific activity of gonadotropin products, both types of products, urinary and recombinant gonadotropins, are on the market at the present (3,4). From a clinical perspective, the decision regarding what kind of gonadotropin to give to a woman undergoing a treatment for fertility is still challenging. One of the points to consider is the possible differences in effectiveness among the different gonadotropins.

<u>rhFSH-alfa reference product and hMG</u>

Gonadotropins extracted from urine and produced from recombinant processes present significant differences in terms of glycosylation profile, purity, consistency and specific activity (4–7).

Randomized clinical trials (RCTs) comparing rhFSH-alfa reference product and hMG on relevant outcomes for patients and physicians are scarce with different results pointing in different directions. A small number of RCTs found no difference in implantation (8,9), pregnancy outcomes (ongoing pregnancy, clinical pregnancy and live birth) (8–10) and pregnancy loss rates per started cycle (8,9). Others have showed that, although the pregnancy outcomes did not differ between the two types of treatment, treatments with rhFSH-alfa reference product led to a significantly higher number of recovered oocytes compared to treatments with hMG (mean (± standard deviation [SD]): 12.29 (±7.80) vs. 9.67 (±5.92); $p<0.001$) (11), (mean (±SD): 14.4 (±8.1) vs. 11.3(±6.0); $p=0.001$) (8) and (mean (±SD): 11.8 (±5.7) vs. 10.0 (±5.4); $p<0.001$) (9) in fresh embryo transfer cycles.

Recently, a non-interventional study (NIS) using real world data (RWD) from 71 in vitro fertilization (IVF) centers in Germany assessed the effectiveness of the two main brands in Germany, rhFSH-alfa reference product (n=17,725 women) vs. Menogon HP, an hMG product (n=10,916 women), over a total of 38,234 cycles (12). In the analysis per patient, data showed better effectiveness outcomes for patients treated with rhFSH-alfa reference product than for patients treated with Menogon HP, both overall and in the GnRH agonist sub-group (the most frequently used protocol in the study period, 2007-2012) after stratification on the GnRH protocol,

and after adjustment on the main confounders. Results showed significantly higher live birth in (1) the total population (adjusted hazard ratio (HR)=1.10, 95% confidence interval (CI): 1.04-1.16) and with GnRH agonists (adjusted HR=1.13, 95% CI: 1.07-1.19); (2) higher ongoing pregnancy in the total population (adjusted HR=1.10, 95% CI: 1.04-1.16) and with GnRH agonists (adjusted HR=1.13, 95% CI: 1.07-1.19); and (3) higher clinical pregnancy in the total population (adjusted HR=1.10, 95% CI: 1.05-1.14) and with GnRH agonists (adjusted HR=1.12, 95% CI: 1.07-1.17)). Results remained significant for all the 3 outcomes listed above in the analysis per first cycle and completed cycles. This supports the hypothesis that differences observed at the purity, consistency and molecule level in respect to glycosylation and other biological characteristics may have also important clinical implications, especially in the context of the GnRH agonist protocol.

rhFSH-alfa reference product and rhFSH-alfa biosimilars

rhFSH-alfa biosimilar preparations have been approved in Europe, as well as in other countries such as, but not limited to, Russian, India, China, Korea, and Argentina. Two rhFSH-alfa biosimilar products are now available on the European Union (EU) market based on the European Medicines Agency's (EMA) guideline for biosimilars approval (Ovaleap®, authorized since 2013; Bemfola®, authorized since 2014) (4).

Approval of biosimilars has the main objective to guarantee treatment accessibility to all the patients in all the markets, in an ideal case, due to lower price of the product with the same safety and efficacy as the reference product (13).

In Europe, according to the guideline of EMA, the Marketing Authorization application dossier of a biosimilar medicinal product shall provide a full quality dossier together with data demonstrating comparability with the reference medicinal product by using appropriate physical-chemical and in vitro biological tests, non-clinical studies and clinical studies (14). Until now, biosimilarity of new rhFSH-alfa preparations have been demonstrated with Phase III clinical trials powered to detect non-inferiority on the number of oocytes retrieved when compared with the reference preparations.

Indeed, according to the EMA guidelines, oocyte number is the most appropriate endpoint to demonstrate the clinical comparability in terms of follicular development of rhFSH-alfa biosimilar and reference preparations. However, live birth, ongoing pregnancy and ovarian hyperstimulation syndrome (OHSS) are considered the most relevant clinical and safety outcomes of ART treatment according to the Guideline for Ovarian Stimulation in IVF or intracytoplasmic sperm injection (ICSI) by the European Society of Human Reproduction and Embryology (ESHRE) (15) and revised glossary of ART terminology by the International Committee for Monitoring Assisted Reproductive Technologies (ICMART) (16). These clinical and safety outcomes were evaluated as secondary endpoints in the submission trials, and the studies were not individually powered to enable comparison of these parameters.

In addition to individual RCTs, a meta-analysis (MA) on rhFSH-alfa reference product vs. rhFSH-alfa biosimilars evaluated live birth as the primary endpoint, and clinical pregnancy and OHSS as secondary endpoints, after the first stimulation cycle (17). This MA confirmed the limited availability of published RCTs that could be included in the analysis. Overall, live birth and CPR were significantly lower in women treated with rhFSH-alfa biosimilars compared to rhFSH-alfa reference product (relative risk (RR) 0.82, 95% CI: 0.69-0.96; and RR 0.82, 95% CI: 0.70-0.95,

respectively). The rate of OHSS was non-significantly increased with rhFSH-alfa biosimilar treatment (RR 1.31, 95% CI: 0.67-2.56).

Protein characterization data available suggested that main differences between the reference (GONAL-f) and biosimilar rhFSH-alfa molecular structure that may impact the pregnancy outcomes are related to protein glycosylation pattern (18).

Outcomes in fertility studies

Most of the current published evidence in fertility treatment strategy and efficacy has been obtained by assessing clinical outcomes in the first cycle with fresh embryo transfer in an RCT setting, and only rarely in a cumulative way, including all fresh and frozen embryo transfers (originated from the same stimulation cycle). However, from a clinical point of view, the cumulative pregnancy and live birth rates are considered the most clinically relevant outcomes in ART, even though very limited data are available due to the typically scarce follow-up data collected at the IVF centers. Cumulative ongoing pregnancy rate (COPR) is considered as a valid predictor for the clinical outcome of cumulative live birth rate (CLBR), resulting from one initiated or aspirated ART cycle (including the cycle when fresh embryos are transferred and subsequent frozen/thawed ART cycles). This is calculated per patient in respect to a single type of treatment received or per started cycle (16).

Time-to-pregnancy (TTP) (or live birth – TTLB) has been lately considered as a relevant clinical outcome of ART. The ICMART defines TTP as the time taken to establish a pregnancy, measured in months or in numbers of menstrual cycles (16). Currently, there is no standard use, or definition, of a time-related outcome measure for fertility-related clinical studies. It depends on the starting point considered, if calculated per patient in respect to a single type of treatment received or per started cycle or any other relevant ART setting, and on considered outcomes (OPR, LBR, COPR, CLBR).

As stated above, usually live birth is reported in pivotal clinical trials (CTs) as a secondary outcome and other meaningful clinical outcomes, including ongoing pregnancy and clinical pregnancy, are usually only partially reported in CTs. Moreover, there is scarcity of data in terms of cumulative pregnancy or live birth across all cycles. The overarching reason for this research gap is the requirement of an extensive study follow-up period to evaluate such endpoints. Therefore, use of real-world data (RWD) can be an efficient way to address these types of outcomes, as patient data are collected longitudinally, often over a long period of time, in routine practice. Moreover, the clinical evidence generated based on RWD (so-called real-world evidence [RWE]) and the associated cost of treatment, is valuable for health-economic evaluations. RWE provides unique evidence for healthcare decision makers from different sectors (private or government) to make their decisions, which could benefit patient access and reduce healthcare burden (19).

Study rationale

The hypothesis tested in this study is that there are possible differences in effectiveness and safety outcomes between rhFSH-alfa reference product and urinary gonadotropins or rhFSH-alfa biosimilar products, with a beneficial effect of rhFSH-alfa reference product in terms of meaningful clinical outcomes.

With the long-term market use of rhFSH-alfa reference and urinary products, and the increased use of rhFSH-alfa biosimilar products, it is important to assess this hypothesis in the routine clinical practice setting. This will be achieved by comparing the effectiveness, and safety, of rhFSH-alfa reference product vs. highly purified (HP-)hMG or rhFSH-alfa biosimilar products, in different GnRH down-regulation regimens, regarding (cumulative) live birth and other effectiveness and safety outcomes that would be relevant for physicians, and most importantly for patients.

8 Research Question and Objectives

This study hypothesizes that controlled ovarian stimulation (COS) with rhFSH-alfa reference product is superior to COS with HP-hMG or rhFSH-alfa biosimilar products with respect to effectiveness and safety outcomes among women undergoing IVF/ICSI treatment in routine clinical practice.

Specifically, the research question is whether rhFSH-alfa reference product is associated with better treatment results than HP-hMG or rhFSH-alfa biosimilar products in terms of pregnancy outcomes (live birth, ongoing pregnancy, clinical pregnancy) per stimulation cycle and per patient [for cumulative clinical outcomes including all cycles with the same drug with fresh and frozen embryo transfers (originated from the same stimulation cycle)]. The study also investigates other clinical effectiveness outcomes (e.g., number of oocytes retrieved) and safety outcomes (e.g., pregnancy loss, OHSS). Lastly, costs associated with IVF/ICSI treatment are described for each type of gonadotropin product. Answering these comprehensive questions will help establish whether rhFSH-alfa reference product should be preferred over the comparator drugs in routine clinical practice.

8.1 Primary Objective

The primary objective of this study is to compare rhFSH-alfa reference product with HP-hMG or rhFSH-alfa biosimilar products regarding the following live birth outcome measures:

a) LBR per initiated IVF/ICSI stimulation cycle

b) CLBR per initiated IVF/ICSI stimulation cycle

c) CLBR in up to five initiated IVF/ICSI stimulation cycles [termed multiple-cycle (MC-)CLBR]

8.2 Secondary Objectives

There are multiple secondary objectives, where secondary objectives 1-4 examine treatment effectiveness and secondary objectives 5-8 treatment safety. Hence, the secondary objectives compare rhFSH-alfa reference product with HP-hMG or rhFSH-alfa biosimilar products regarding the following outcomes:

1. a) CPR and OPR per initiated IVF/ICSI stimulation cycle, and b) cumulative clinical pregnancy rate (CCPR) and COPR per initiated IVF/ICSI stimulation cycle

2. Number of oocytes retrieved, embryos transferred, embryos cryopreserved, and utilizable embryos per initiated IVF/ICSI stimulation cycle and per oocyte retrieval cycle

3. Implantation rate

4. TTLB

5. Rate of pregnancy loss for the first (and successive) IVF/ICSI stimulation cycle(s)

6. Rate of multiple pregnancy for the first (and successive) IVF/ICSI stimulation cycle(s)

7. Number of cancelled cycles for the first (and successive) IVF/ICSI stimulation cycle(s)

8. Number of OHSS cases for the first (and successive) IVF/ICSI stimulation cycle(s)

An additional secondary objective describes treatment-associated costs. Hence, the secondary objective 9 is to describe the following outcomes for the rhFSH-alfa reference product, HP-hMG and rhFSH-alfa biosimilar products individually and overall, across all products:

9. Costs associated with IVF/ICSI treatment (including drugs and other treatments), miscarriage/birth, and adverse events (i.e., OHSS)

9 Research Method

9.1 Study Design

9.1.1 Design Overview

This is a non-interventional study based on secondary data from national population-based registers in Denmark and Sweden. The study uses a cohort design and is conducted as a comparative effectiveness and safety study with head-to-head comparisons of drugs used for treatment of infertility and in ART. The study drugs are rhFSH-alfa reference product (drug of interest), HP-hMG and rhFSH-alfa biosimilar products (comparator drugs). Together with the rhFSH-alfa reference product, the comparator drugs constitute the main gonadotropins used for COS in ART in both countries of study in recent years. In 2019, the three study drugs comprised 96% of all dispensed prescriptions for gonadotropins in Sweden. In Denmark, the study drugs comprised 89% of all volume sales in the same year[1]. Comparisons will be made between the rhFSH-alfa reference product and each comparator drug with respect to clinical effectiveness and safety. Treatment-associated costs will be described. Thus, the newer generation rhFSH-alfa reference product will be evaluated both against the older generation products, HP-hMG, and the newest generation rhFSH-alfa products, the biosimilars.

[1] Publicly available information from each country's online statistical database on pharmaceuticals (*Denmark:* https://medstat.dk/en; *Sweden:* https://sdb.socialstyrelsen.se/if_lak/val_eng.aspx). The following ATC codes were used in the search for statistics on sales volumes and number of dispensed prescriptions: G03GA02 (hMG), G03GA04 (urofollitropin), G03GA05 (follitropin alfa), G03GA06 (follitropin beta), G03GA09 (corifollitropin alfa), G03GA30 (combinations). Accessed 10 February 2021.

Data will be obtained from nationwide, electronically recorded, population-based registers with prospective data collection. In particular, the IVF register in each country will be used to establish the study population. The IVF registers hold comprehensive information on ART treatments conducted at all fertility clinics, publicly- and privately-operated, in Denmark and Sweden, respectively. The study will therefore enable inclusion of information on IVF treatments among the entire female population of each country. The unique personal identification number assigned to individuals at birth or immigration in each country further allows for linkage of individual-level data across multiple registers and subsequently the creation of country-specific customized study databases. Data will be obtained on IVF/ICSI treatment characteristics, dispensed prescription drugs (including exposure to the study drugs, GnRH protocol used, etc.), pregnancy and other treatment outcomes, medical history, laboratory test results (Denmark only), and costs. The contents, set-up and structure of each national register is similar across the two countries, including coding classification systems such as the International Classification of Diseases, 10th revision (ICD-10) and the Anatomical Therapeutic Chemical (ATC) classification system. The same type of analyses can therefore be performed in both countries, with little need for adaptation. Individual-level data will be analyzed separately for each country. Meta-analysis may be used, as appropriate, to provide summary estimates for the entire study population across both countries.

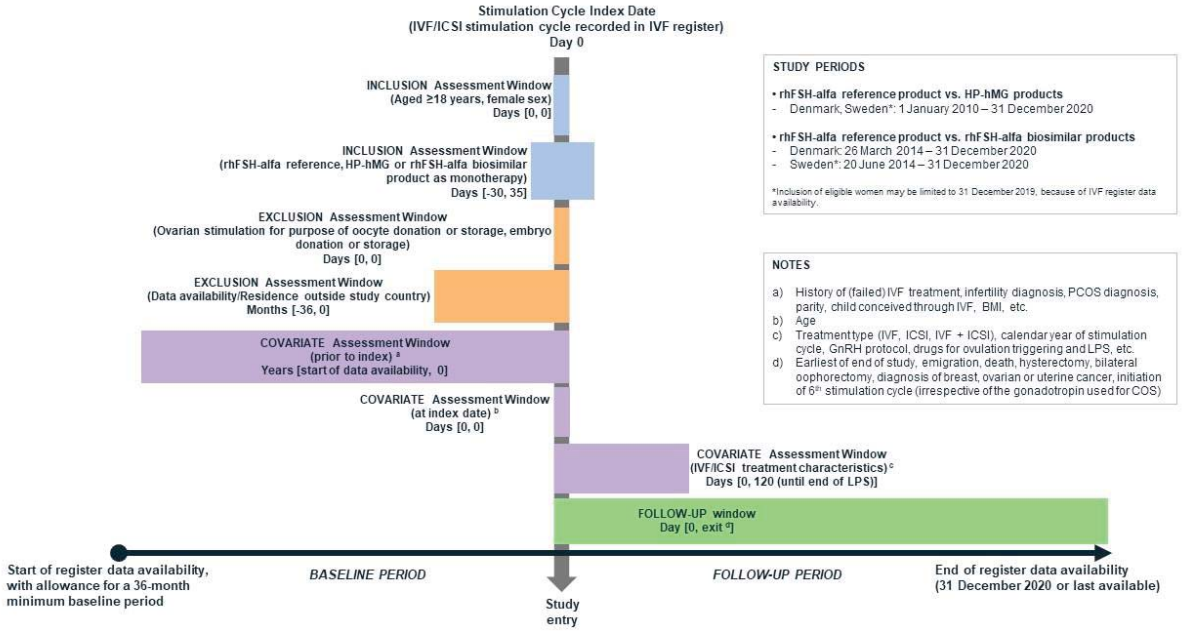
The study will include IVF or ICSI stimulation cycles and the FET cycles linked to the same stimulation cycles among women who received rhFSH-alfa reference product, HP-hMG or rhFSH-alfa biosimilar products for COS (Figure 1). The overall qualifying event for study inclusion will thus be a recorded IVF/ICSI stimulation cycle with one of the study drugs. Three study cohorts will be created, including stimulation cycles with rhFSH-alfa reference product, HP-hMG, and rhFSH-alfa biosimilar products, respectively. Women may contribute cycles to more than one study cohort. Any linked FET cycles will be assigned to the same cohort as the stimulation cycle it originated from. The study will be restricted to the first five stimulation cycles observed in each woman in the IVF registers, regardless if they occurred in the study period and the gonadotropin (study drug or non-study drug) used for COS. The study is restricted to the first five stimulation cycles as it is known that the CLBR per cycle decreases over the first five cycles and is maintained at a similarly low level beyond the fifth stimulation cycle and the number of women continuing treatment diminishes substantially with each advancing cycle number (20,21). Further, the study population will be restricted to women who, on the day of treatment initiation (ovarian stimulation start) and 3 years prior, resided in the country where the treatment was performed. This will allow for a minimum baseline period of 3 years during which any previous IVF/ICSI treatment and information on potential confounders can be assessed.

The overall study period will be between 2010 and 2020. Different study period starts will be applied depending on which two drugs are being compared. For the comparison between rhFSH-alfa reference product and HP-hMG, the study period will start 1 January 2010 for both countries. This date is selected to ensure the study is sufficiently powered while also reducing heterogeneity in patient management over time. Changes that have occurred over time include moving towards vitrification rather than slow freezing for embryo cryopreservation and hence also moving from cleavage stage to blastocyst embryo transfers, elective single embryo transfers rather than double embryo transfers, and antagonist protocols rather than agonist protocols. For the comparison between rhFSH-alfa reference product and rhFSH-alfa biosimilar products, the study period will start 26 March 2014 for Denmark and 20 June 2014 for Sweden. These are the earliest dates when a rhFSH-alfa biosimilar product (Bemfola) received reimbursement status in each country. For

Sweden, where the IVF register is updated annually and has a lag time of approximately two years (see Section 9.4.2), inclusion of women undergoing IVF/ICSI treatment may end on 31 December 2019, depending on data availability at the time of data extraction. No such general restriction will be applied for Denmark, as the IVF register is updated monthly. In terms of follow-up of IVF/ICSI treatment and pregnancy outcomes, the study will end on 31 December 2020 (or last available date) for each country. Specifically, for the main analyses of the primary objective (LBR, CLBR and MC-CLBR) and secondary objective 4 (TTLB), a minimum follow-up period of 1 year has been chosen to allow sufficient time to achieve (at least) one live birth and a reasonable number of complete cycles for reliable estimates of live birth rates. Hence, the corresponding analyses are restricted to cycles initiated on or before 31 December 2019 (Sets 1 and 5 in Table 5). The minimum follow-up period will be extended to 2 years in a sensitivity analysis to explore the potential effect this has on the estimated MC-CLBR and TTLB (see sensitivity analysis 3 in Table 8).

To reduce possible bias linked to confounding by baseline characteristics, inverse probability of treatment weighting (IPTW), based on propensity scores (PS) (22), will be used to balance the distribution of potential confounders across the study cohorts.

Figure 1. Study design



BMI: body mass index; COS: controlled ovarian stimulation; GnRH: gonadotropin-releasing hormone; HP-hMG: highly purified human menopausal gonadotropin; ICSI: intracytoplasmic sperm injection; IVF: in vitro fertilization; LPS: luteal phase support; PCOS: polycystic ovarian syndrome; rhFSH: recombinant human follicle-stimulating hormone.

9.1.2 Outcomes

The primary and secondary outcomes detailed below will be identified using linked data from the IVF registers, patient registers and medical birth registers in Denmark and Sweden, as needed. An overview of how the registers will be used for outcome identification is provided in Appendix 3

(Section 14.3.1), along with diagnosis and procedure codes to ascertain the outcomes (where applicable).

9.1.2.1 Primary

The primary outcome of this study is live birth. It is defined generally as the complete expulsion or extraction from a woman of a product of fertilization, after 22 completed weeks of gestational age; which, after such separation, breathes or shows any other evidence of life, such as heartbeat, umbilical cord pulsation or definite movement of voluntary muscles, irrespective of whether the umbilical cord has been cut or the placenta still attached or not (16).

Information on live birth will be obtained from the medical birth register in each country and the IVF register in Sweden. Only live births that can be linked to a given IVF/ICSI stimulation cycle included in the study will be considered. The live birth of multiples will be counted as one live birth in this study.

For the assessment of the primary objective and secondary objective 4, the different measures of live birth are defined as follows:

- LBR is the probability of a live birth after one initiated stimulation cycle, considering fresh embryo transfer.
- CLBR is the number of live births per initiated stimulation cycle, considering all fresh and subsequent frozen embryo transfers linked to that stimulation cycle and counting all live births from the same transfers.
- MC-CLBR is the probability of a live birth in up to five stimulation cycles, considering all fresh and subsequent frozen embryo transfers linked to each of those stimulation cycles, until the first live birth occurs.
- TTLB is the time (in number of cycles) until live birth in up to five stimulation cycles, considering all fresh and subsequent frozen embryo transfers linked to each of those stimulation cycles, until the first live birth occurs.

For the calculation of the above measures, refer to Section 9.7.3.3 (Table 7).

9.1.2.2 Secondary

There are several secondary outcomes, which are connected to the secondary study objectives:

1. Clinical pregnancy is defined generally as pregnancy diagnosed by ultrasonographic visualization of one or more gestational sacs. It includes intrauterine pregnancy and clinically documented ectopic pregnancy (16). In this study, multiple embryos will be counted as one clinical pregnancy.

   Intrauterine pregnancy will be ascertained in the IVF registers. Diagnosis and procedure codes used to ascertain ectopic pregnancy in the national patient registers are provided in Appendix 3 (Section 14.3.1).

   For the assessment of secondary objective 1, clinical pregnancy is measured as CPR and CCPR, which are defined analogously to LBR and CLBR:

- CPR is the probability of a clinical pregnancy after one initiated stimulation cycle, considering fresh embryo transfer.

- CCPR is the number of clinical pregnancies per initiated stimulation cycle, considering all fresh and frozen embryo transfers lined to that stimulation cycle and counting all clinical pregnancies from the same transfers.

2. Ongoing pregnancy is defined in this study as pregnancy still ongoing at 22 completed weeks of gestational age. It will be ascertained based on a record of miscarriage or induced abortion (before 22 completed weeks of gestational age), or gestational age at delivery in the national patient registers or the IVF register in Sweden. Multiples will be counted as one ongoing pregnancy.

   For the assessment of secondary objective 1, ongoing pregnancy is measured as OPR and COPR, which are defined analogously to LBR and CLBR:

   - OPR is the probability of an ongoing pregnancy after one initiated stimulation cycle, considering fresh embryo transfer.

   - COPR is the number of ongoing pregnancies per initiated stimulation cycle, considering all fresh and frozen embryo transfers linked to that stimulation cycle and counting all ongoing pregnancies from the same transfers.

   Only clinical pregnancies and ongoing pregnancies that can be linked to a given IVF/ICSI stimulation cycle included in the study will be considered.

3. Oocytes retrieved is defined generally as the total number of oocytes retrieved from an ovarian follicular aspiration following COS. Information on the number of oocytes retrieved will be ascertained in the IVF registers.

4. Embryos transferred is defined generally as the total number of embryos placed into the uterus following IVF or ICSI. Information on the number of embryos transferred will be ascertained in the IVF registers. This outcome is applicable to both fresh and frozen embryo transfers.

5. Embryos cryopreserved is defined generally as the total number of embryos cryopreserved following COS, follicular aspiration, and embryo cultivation, using the process of slow freezing or vitrification to preserve cleavage-stage embryos. Information on the number of embryos cryopreserved will be ascertained in the IVF registers.

6. Utilizable embryos are defined as the sum of number embryos transferred and number embryos cryopreserved.

7. Implantation is defined generally as the attachment and subsequent penetration by a zona-free blastocyst into the endometrium. This process starts 5 to 7 days after fertilization of the oocyte usually resulting in the formation of a gestation sac (16). It will be ascertained in the IVF registers based on ultrasonographic visualization of the number of gestational sacs.

8. Pregnancy loss is defined generally as the miscarriage or induced abortion of an intrauterine pregnancy before 22 completed weeks of gestational age. Diagnosis and procedure codes used to ascertain pregnancy loss are provided in Appendix 3 (Section 14.3.1). Only complete pregnancy losses, *i.e.*, involving all fetuses in case of multiple pregnancy, will be considered.

9. Multiple pregnancy is defined generally as a pregnancy with more than one embryo or fetus. It will be ascertained in this study based on ultrasonographic visualization of two or more gestational sacs and information on pregnancy outcome, to account for cases of monozygotic twins. Because monozygotic twins share the same gestational sac, an algorithm will be developed which considers data on intrauterine pregnancy loss and data on live or stillbirth. In this way, reduction of multiple pregnancy is accounted for, which is particularly important for Denmark as the Danish In-Vitro Fertilization Register (DIVF) does not record pregnancy outcome. The algorithm will be detailed in the statistical analysis plan (SAP).

10. Cycle cancellation is defined generally as a cycle in which ovarian stimulation or monitoring has been initiated with the intention to treat, but which did not proceed to follicular aspiration or, in the case of a thawed or warmed embryo, did not proceed to embryo transfer.
    Information on cycle cancellation will be ascertained in the IVF registers.

11. Ovarian hyperstimulation syndrome is defined generally as an exaggerated systemic response to COS characterized by a wide spectrum of clinical and laboratory manifestations. It will be ascertained in this study based on a recorded primary or secondary diagnosis of OHSS within inpatient or outpatient specialized care within 70 days (10 weeks) after oocyte retrieval or as recorded in the IVF registers in a given IVF/ICSI stimulation cycle. For descriptive purposes, OHSS will be classified as early if it occurred within 9 days after oocyte retrieval, and late if it occurred 10 or more days after oocyte retrieval. The day of hospital admission (inpatient care) or hospital visit (outpatient care) will be defined as the onset of OHSS (23). A woman may experience both early and late OHSS. Thus, for the description of OHSS, a woman may be counted towards the early and late classification in a given cycle. For the combined measure to be used in the comparative analyses, only one occurrence of OHSS per cycle will be counted. If a new cycle was initiated within 70 days after oocyte retrieval, the previous stimulation cycle will be excluded from further descriptive and comparative analyses with respect to OHSS from the date of treatment start for the new cycle. The diagnosis code used to ascertain OHSS is provided in Appendix 3 (Section 14.3.1).

12. Treatment-associated costs are defined as direct medical costs of treatments (drugs and other treatments), pregnancy loss/birth, and adverse events (i.e., OHSS) resulting from an IVF/ICSI stimulation cycle.

Detailed definitions for ascertainment of the secondary outcomes using ICD-10 and procedure codes, as applicable, are provided in Appendix 3 (Section 14.3.1). For the calculation of the secondary outcome measures, refer to Section 9.7.3.3 (Table 7).

9.2 Setting

9.2.1 Study Population

This study will comprise IVF/ICSI stimulation cycles and FET cycles linked to the same stimulation cycles among women who received rhFSH-alfa reference product, HP-hMG or rhFSH-alfa biosimilar products for COS. The countries of study are Denmark and Sweden. The study population will be identified using registers with comprehensive nationwide coverage. Hence, all eligible IVF/ICSI cycles among the entire female population in each country will be included and the results from the study are expected to be representative of that population.

Women aged 18 years or more at ovarian stimulation start during the study period will be identified as eligible for inclusion in the study from each country's IVF Register. Individual-level data will then be linked to the Prescribed Drug Register, the National Patient Register, and the Medical Birth Register to further determine eligibility and ascertain baseline characteristics and safety and pregnancy outcomes. More details on each register are provided in Section 9.4.

For inclusion, all stimulation cycles will be assessed on a set of inclusion and exclusion criteria. The overall study population will include all women with at least one IVF/ICSI stimulation cycle fulfilling all eligibility criteria. Analysis-specific selection criteria, depending on the outcome of interest, will also be applied (see Section 9.7.1).

For inclusion of a stimulation cycle in the study, women must fulfill all of the following inclusion criteria:

1. *Initiated IVF/ICSI stimulation cycle with rhFSH-alfa reference product, HP-hMG or rhFSH-alfa biosimilar product monotherapy for COS during the study period*[2]

2. *Aged 18 years or more at stimulation cycle index date*

3. *Female sex at stimulation cycle index date*

Stimulation cycles are not eligible for this study if any of the following exclusion criteria are fulfilled:

1. *History of 5 or more IVF/ICSI stimulation cycles prior to the stimulation cycle index date*

2. *Ovarian stimulation for the purpose of oocyte donation, oocyte storage, embryo donation, oncological or other medically indicated embryo storage*[3]*, or preimplantation genetic testing (PGT)*

3. *Non-availability of individual-level data on IVF/ICSI treatment, dispensed drugs and medical history for 36 months or more prior to and including the stimulation cycle index*

[2] For Sweden, inclusion of eligible women may be limited to 31 December 2019 because of Q-IVF data availability.
[3] For Denmark, oncological/medically indicated embryo storage will be derived based on information on cancer and other relevant diagnoses and associated treatments recorded in the DIVF and other national registers. Relevant codes will be detailed in the SAP. For Sweden, this information is directly available in the Q-IVF.

date, operationalized as the woman not being resident of the country where the IVF/ICSI treatment was performed during the 36-month period

Study time periods and stimulation cycle index date are defined in Section 9.2.3.

9.2.2 Definition of Study Cohorts and Description of Treatments

Among the women included in the study, three study cohorts will be created at the level of the IVF/ICSI stimulation cycle based on which drug the woman received for COS: rhFSH-alfa reference product, HP-hMG or rhFSH-alfa biosimilar product. Only one type of gonadotropin use will be allowed for each stimulation cycle. Hence, cycles where mixed gonadotropin use was observed will not be included in the study. Mixed gonadotropin use may occur for example if, at prescription refill at the pharmacy, the drug (brand) on the written prescription was exchanged for an equivalent drug (same active substance, strength, formulation; therapeutically equipotent).

A woman may contribute maximum five stimulation cycles to the same cohort if she had several cycles with the same study drug. She may contribute cycles to more than one study cohort if she had distinct stimulation cycles with different study drugs. Between stimulation cycles with a study drug, a woman may have received a non-study gonadotropin for COS. History of IVF/ICSI treatment, including with a non-study gonadotropin, will be controlled for in the analysis (in the PS estimation). Overall, no more than five stimulation cycles and only up to the fifth stimulation cycle will be considered in this study, regardless if all cycles were observed in the study period and the type of gonadotropin used for COS.

FET cycles originated from an IVF/ICSI stimulation cycle included in the study will be assigned to the same cohort as that stimulation cycle. FET cycles originated from a stimulation cycle in which one of the study drugs was used but which was not included in the study (e.g., because of PGT), will not be considered. Different methods for cycle linkage will be used for Denmark and Sweden. For Sweden, the IVF register includes information enabling an FET cycle to be linked to the stimulation cycle it originated from. This information is the follicular aspiration date, which is recorded for IVF/ICSI stimulation cycles and FET cycles alike. For Denmark, no such direct link exists in the IVF register.

The current (24) and previous (25,26) Danish guidelines on ART state that all frozen embryos should be used before proceeding to the next stimulation cycle. However, in practice, before approximately 2015, many clinics allowed for one or two cryopreserved embryos to remain in the freezer while the woman started a new stimulation cycle (PPD , personal communication by email, PPD PPD). This practice will impact the possibility of linking FET cycles to the appropriate stimulation cycle and thus assigning FET cycles to the correct study cohort. It may also impact the estimation of CLBR, MC-CLBR, TTLB and all other outcomes including FET cycles (see Section 9.1.2). Where an FET cycle was preceded by multiple stimulation cycles, of which more than one resulted in cryopreserved embryos, different assumptions will be made depending on which type of gonadotropin was used for COS. If all preceding stimulation cycles used the same gonadotropin, it will be assumed that the FET cycle is linked to the most recent stimulation cycle. If the preceding stimulation cycles used different gonadotropins, the FET cycle will be excluded from all analyses. This latter approach is not applicable to the estimation of TTLB, as a woman will be censored from that analysis at switch to

a different gonadotropin (see Section 9.7.3.5). The assumptions will be tested in a sensitivity analysis of the primary study objective (see Analysis 5 in Table 8). In instances where an FET cycle was preceded by one stimulation cycle only or multiple stimulation cycles of which only one resulted in one or more cryopreserved embryos, that FET cycle and the embryo(s) used in it can be reliably linked to the stimulation cycle it originated from. Alternative approaches for FET cycle linkage will be considered if the methods suggested for the main and sensitivity analyses are judged inadequate.

Lastly, cycles are required to be initiated, but not completed (fresh or frozen embryos transferred), to be included in a study cohort. As it is not available from the national registers, gonadotropin starting dose will not be considered when creating the study cohorts.

9.2.3 Study Time Periods

The overall **study period** is from 2010 through 2020. Different start dates are defined based on which two drugs are compared and, for the comparison with rhFSH-alfa biosimilar products, which country is analyzed. The study period ends on 31 December 2020 (or last available data) for both study drug comparisons and countries. Homogeneity in patient management over time, market entry of the comparator drugs and register lag times determine the start and end dates of the study period:

- rhFSH-alfa reference product vs. HP-hMG:
    - Denmark, Sweden: 1 January 2010 – 31 December 2020
- rhFSH-alfa reference product vs. rhFSH-alfa biosimilar products:
    - Denmark: 26 March 2014 – 31 December 2020
    - Sweden: 20 June 2014 – 31 December 2020

For Sweden, inclusion of eligible women may need to be restricted to 31 December 2019, considering study timelines (see Section 6) and the typical two-year lag time for data availability from the Swedish National Quality Registry of Assisted Reproduction (Q-IVF) (see Section 9.4.2).

The overall **index date** is the date when the woman's *first* eligible IVF/ICSI stimulation cycle started (defined as start of ovarian stimulation) in the study period. The **stimulation cycle index date** is the date when each stimulation cycle started. For FET cycles linked to an included stimulation cycle, the cycle start date is the Last Menstrual Period (LMP) date for natural FET cycles or the date hormonal stimulation (replacement) treatment was started for stimulated FET cycles, as recorded in each country's IVF register.

The **baseline period** is defined as the 36-month minimum period prior to and including the (stimulation cycle) index date. During this period, (cycle) eligibility will be determined, and baseline covariates assessed. The baseline period may include time before the age of 18 years, for women who had IVF/ICSI treatment specifically at the age of 18 years. All available data from the registers will be used to obtain information on important covariates. Baseline data from the national patient registers, however, will be limited to the period in which ICD-10 coding has been

used comprehensively in each country (starting from 1994 in Denmark and 1998 in Sweden). The exact length of the baseline period will therefore vary between cycles within the same patient.

The overall **follow-up period** among women included in the study is the time between the index date and the exit date. The exit date is the censoring date, as defined according to the following general censoring criteria: initiation of a sixth stimulation cycle (regardless of gonadotropin used for COS), end of the study period, emigration (available for Denmark only), death, hysterectomy, bilateral oophorectomy or diagnosis of breast, ovarian or uterine cancer, whichever occurs first. Moreover, each included stimulation cycle will be followed from the stimulation cycle index date until the date when each outcome is assessed; this date varies between the analyses of different objectives and outcomes (see Section 9.7). For a graphical illustration of the study time periods, see Figure 1. As applicable, diagnosis and procedure codes used to define the censoring criteria are provided in Appendix 3 (Section 14.3.2).

9.3 Variables

Derived and transformed data needed for the analysis are described in Section 9.7.2.

9.3.1 Exposure Variables

Exposure of an IVF/ICSI stimulation cycle to rhFSH-alfa reference product (follitropin alfa, ATC code G03GA05), HP-hMG (ATC code G03GA02) or rhFSH-alfa biosimilar product (ATC code G03GA05) will be established based on information from the IVF registers and, primarily, prescriptions dispensed in community pharmacies. While the IVF registers indicate which drug (international nonproprietary name (INN)/ATC) was used for ovarian stimulation, the prescribed drug registers comprise more detail. Additionally, in Sweden, drug information is expected to be available from the Q-IVF starting only from 2019.

Hence, information on the dispensed prescription will include substance name (INN), brand name, ATC code, strength of drug, dosage form, pack size, number of dispensed packages and dispensed amount of drug [in Defined Daily Doses (DDD)]. Using the brand name recorded with the dispensed prescription, it will be possible to distinguish between rhFSH-alfa reference product and rhFSH-alfa biosimilar product. Because information on prescribed daily dose is expected to be incomplete and/or recorded in free-text format, it will not be considered for estimation of total gonadotropin dose.

Using data from the prescribed drug registers, a time window of 30 days before and 35 days after the stimulation cycle index date (27) will be defined to ascertain that gonadotropin monotherapy with a study drug was used for COS. The impact of using a narrower time window for exposure ascertainment will be explored in a sensitivity analysis (see Analysis 1 in Table 8). Any FET cycle(s) linked to an IVF/ICSI stimulation cycle included in the study will be considered exposed to the same study drug as the cycle it originated from.

Total gonadotropin dose will not be considered in the analyses of the comparative study objectives (1-8). However, conditional that the data from the Q-IVF is comprehensive, total gonadotropin dose may be considered for the estimation of drug cost in the descriptive cost analysis for Sweden (see Section 9.7.3.7), albeit for a limited portion of the study period.

9.3.2 Other Variables (Covariates)

In addition to the outcome and exposure variables outlined in Section 9.1.2 and Section 9.3.1, the covariates of interest in this study are listed below and classified according to their intended use in the analysis (Table 1).

The variables are generally described as measured in relation to each IVF/ICSI stimulation cycle; however, variables used to derive the outcome variables may also apply to any linked FET cycles, depending on the outcome under study. Variables indicated as being candidates for regression modeling are determined after the decision on which gonadotropin to use for COS has been made. These variables will therefore not be included in the PS but may instead be added as independent covariates in the IPTW-weighted regression model (see Section 9.7.3.1 for further details).

For variables identified in the registers by ATC, ICD-10 and/or procedure codes, these operational definitions are provided in Appendix 3 (Section 14.3.3).

Table 1. List of covariates and their use in the study

Variable	Measurement scale	Baseline characteristic	Candidate variable for PS	Candidate variable for additional adjustment
Demographic characteristics				
Age at stimulation cycle index date	Continuous	X	X	
IVF/ICSI treatment characteristics				
Treatment type (ICSI, IVF, ICSI + IVF; FET)	Categorical	X	X	
Type of clinic (private, public)	Categorical	X	X	
Calendar year of stimulation cycle	Continuous	X		
GnRH protocol (agonist, antagonist)	Categorical	X	X	
Cycle cancellation (timing within cycle)	Categorical	N/A (used to derive outcome variable)		
Date of cycle cancellation	Continuous	N/A (used to derive outcome variable)		
Type of drug for ovulation triggering	Categorical	X		X

Variable	Measurement scale	Baseline characteristic	Candidate variable for PS	Candidate variable for additional adjustment
Type of drug for luteal phase support	Categorical	X		X
Date of follicular aspiration	Continuous	N/A (used to derive outcome variable)		
Date of embryo transfer	Continuous	N/A (used to derive outcome variable)		
Number of days of embryo cultivation[+]	Numerical	X		X
Date of embryo cryopreservation	Continuous	N/A (used to derive outcome variable)		
Embryo cryopreservation method (slow freezing, vitrification)	Categorical	X		X
Date of embryo thawing	Continuous	N/A (used to derive outcome variable)		
Number of embryos thawed	Continuous	N/A (used to derive outcome variable)		
Date of pregnancy test	Continuous	N/A (used to derive outcome variable)		
Pregnancy test result or indicator for pregnancy	Continuous	N/A (used to derive outcome variable)		
Date of ultrasound	Continuous	N/A (used to derive outcome variable)		
Date of pregnancy loss	Continuous	N/A (used to derive outcome variable)		
Other clinical characteristics				
Height/weight or BMI	Continuous	X	X	
Obesity	Categorical	X	X	
Bariatric surgery	Categorical	X	X	
Smoking status[*]	Numerical	X	X	

Variable	Measurement scale	Baseline characteristic	Candidate variable for PS	Candidate variable for additional adjustment
Type 1 diabetes mellitus	Categorical	X	X	
Type 2 diabetes mellitus	Categorical	X	X	
Antidiabetic drugs	Categorical	X	X	
History of thromboembolic events	Categorical	X	X	
Use of thrombosis prophylaxis	Categorical	X	X	
Porphyria and other relevant conditions[#]	Categorical	X	X	
History of any cancer	Categorical	X	X	
Medical history, fertility-related				
Number of previous failed COS	Categorical	X	X	
History of failed IUI	Categorical	X	X	
History of COS with a study drug	Categorical	X	X	
History of COS with a non-study drug	Categorical	X	X	
Previous delivery with at least one live birth following IVF	Categorical	X	X	
Parity	Categorical	X	X	
OHSS	Categorical	X	X	
Amenorrhea or oligomenorrhea	Categorical	X	X	
PCOS	Categorical	X	X	
Endometriosis	Categorical	X	X	

Variable	Measurement scale	Baseline characteristic	Candidate variable for PS	Candidate variable for additional adjustment
Female infertility diagnosis	Categorical	X	X	
Female factor infertility*	Categorical	X	X	
Male factor infertility*	Categorical	X	X	
Idiopathic factor infertility*	Categorical	X	X	
Medical history, laboratory test results (value at most recent record on or prior to stimulation cycle index date)†				
Anti-Müllerian hormone, pmol/L	Continuous	X	X	
FSH, IU/L	Continuous	X	X	
Thyrotropin, mIU/L	Continuous	X	X	
Prolactin, mIU/L	Continuous	X	X	

BMI: body mass index; COS: controlled ovarian stimulation; FET: frozen embryo transfer; FSH: follicle-stimulating hormone; GnRH: gonadotropin-releasing hormone; ICSI: intracytoplasmic sperm injection; IUI: intrauterine insemination; IVF: in vitro fertilization; IU: international unit; L: liter; mIU: milli-international unit; OHSS: ovarian hyperstimulation syndrome; PCOS: polycystic ovarian syndrome; pmol: picomole; PS: propensity score.

Notes:

+Not available in Denmark and available from 2015 in Sweden; can be approximated by subtracting the date of follicular aspiration from the date of embryo transfer for stimulation cycles.

*Available for Denmark only (from the DIVF).

#Full list of conditions to be considered is provided in Appendix 3 (Section 14.3.3).

†Available for Denmark only [from the Danish Register of Laboratory Results for Research (RLRR); see Section 9.4.1]. Because completeness is uncertain, these data will be used in a sensitivity analysis only, as described in analysis 4 in Table 8.

9.4 Data Source

The Danish and Swedish data sources are the property of, maintained and made accessible for research by permission of governmental and regional authorities. The health care system is a tax-funded, one-payer system with coverage for all residents and only little co-payment from the patients.

Since introducing the Danish Civil Registration System in 1968, demographics (age, sex, geographical region), migration and vital statistics data have been registered electronically on a daily basis for all Danish residents. Civil registration in Sweden dates back to the 17th century and was transferred from the local parishes to the state in 1991. Demographic data (age, sex, geographical region), migration and vital statistics data for all Swedish residents can be found in the Total Population Register (held by Statistics Sweden) from 1968 onwards.

Every individual in Denmark and Sweden is provided with a unique personal identification number at birth or upon immigration which allows for follow-up until death or emigration. The personal identification number is used for many administrative purposes, such as an identifier in population and health care registers.

The personal identification number forms the basis for the precise, deterministic linkage of individual-level data between different registers and databases in Denmark and Sweden. Given that the personal identification number is used as a unique identifier for a multitude of purposes, it allows for the creation of a tailored database with individual-level data for any given study.

The Danish and Swedish health care data sources are typically established as registers for other reasons than research (examples: control and funding of hospitals, payment of general practitioners for services provided to patients, control system for patients' co-payment of prescription drugs, and authorities' control of treatments using assisted reproduction techniques). Thus, the registers are a mirror of clinical practice in each country. Furthermore, the data sources are also used extensively for epidemiological research.

The DIVF, the Civil Registration System (CPR), the Danish Medical Birth Register (DMBR), the Danish National Patient Register (DNPR), the Danish Register of Medicinal Product Statistic (RMPS), the Q-IVF, the Swedish Medical Birth Register (SMBR), the Swedish National Patient Register (SNPR), the Swedish Prescribed Drugs Register (SPDR) and Swedish Cause of Death Register are the databases for identification of the study population and extraction of exposure, covariate and outcome variables. Thus, the information needed is assembled by linkage of several health care registers using the unique personal identification number.

9.4.1 Denmark

The Civil Registration System (CPR) holds base information on personal identification number, date of birth, sex, immigration/emigration dates, ethnicity/country of origin, date of death, etc. since 1970.

The Danish In-vitro Fertilization Register (DIVF) was established to provide evidence for the Danish National Board of Health's monitoring of ART treatments in Denmark as wells as providing data for research. The register contains information on ART treatments at both public and private fertility clinics in Denmark, including the cause of the infertility, fertilization method, origin of semen and oocytes, and vitrification methods. Pregnancy outcome is not recorded. The DIVF has been in operation since 1994, with later improvements in 2006. Starting from this year, information was added on, e.g., IUI, the woman’s and partner’s life style (weight, alcohol and smoking habits), results from ultrasound examination, including number of gestational sacs with/without heartbeat, choice of cryopreservation method, and whether transfer of one embryo only (fresh or frozen) was an elective or non-elective procedure. The register is updated monthly.

It is possible to submit changes to already reported information. With a lag of approximately 2 months, it is estimated to capture the majority of subsequent changes to already reported information.

The Danish National Patient Register (DNPR) contains information on the Danish population's encounters at hospitals (secondary care) (28,29). The register is a central tool for the health care authorities to monitor diseases and treatments, and the activities at the individual hospitals. Data on activities and diagnoses in primary care at general practitioners (GPs) and specialists in private practice is not included. The register contains information on date of admission and discharge, type and setting of care (inpatient or outpatient), diagnoses (classified according to the ICD-10 classification), examinations, surgical and non-surgical procedures, and Diagnosis Related Group (DRG) costs. Data is available from 1977 for inpatient somatic admissions. Since 1995, outpatient visits, emergency room visits and encounters at psychiatric wards are also included. DRG costs are available since 2002.

The Danish Medical Birth Register (DMBR) is an enrichment of data already recorded in the DNPR and compiles relevant information of pregnancies that has led to childbirth in Denmark at home or in a hospital (30). It contains information on characteristics of mother and child, pregnancy and delivery characteristics including induction procedures and gestational age, outcomes of pregnancy and delivery, and the mother's history of previous abortions and/or deliveries. The DMBR includes linkage between infant, mother and father. The DMBR has been in operation since 1973, with later improvements, particularly in 1997. The register is updated once a year after end of December with a lag of 13 months, partly due to the enrichment on information on date of death within the first year after birth.

The Danish Register of Medicinal Product Statistics (RMPS) contains patient-level data on all prescription drugs filled by patients at community pharmacies (31,32), but not medication administered at hospitals. The administrative purpose of the register is to administer how the cost is split between the health care system, the municipality and the patient's co-payment. Thus, completeness and accuracy of data collection is very high. The register contains information on the date of purchase, item number, product name, ATC code, strength per unit, quantity of DDDs (as per the World Health Organization (WHO)) per package and number of packages filled. Data is available from 1995 and onwards. Since April 2004, information on medical indication for prescription and daily prescribed dose by physician has also been available, but completeness and validity are affected by a non-compulsory obligation to record this information.

The Danish Register of Laboratory Results for Research (RLRR) is managed by the Danish Health Data Authority. It has information on laboratory tests at the country's major clinical biochemical and clinical immunological laboratories. Laboratory results are first collected in the National Lab Databank. General practitioners, specialists in private practices and private fertility clinics can forward samples for analysis to the hospital laboratories, when they do not have the capacity/equipment to do the analysis themselves, and these tests are also included in the National Lab Databank. All this information is transferred to the RLRR later. If a patient, proactively, has denied consent to exchange lab results, the results are not transferred to the RLRR.

Since 2013, the five regions in Denmark has gradually started transferring data to the Danish National Lab Databank,, and subsequently the RLRR, but some historical data has also been

transferred to the National Lab Databank The National Lab Databank does not cover the whole nation yet, since all laboratories are not affiliated, and completeness is not reported. However, of all laboratory test results coded with NPU codes across the affiliated laboratories, 95% are currently in the National Lab Databank.

Two types of tests are included in the National Lab Databank : a) analysis of blood, urine, joint fluids and spinal fluids for the purpose of preventing, diagnosing and controlling the treatment of human diseases; b) blood type determination and examination of blood in, e.g., pregnancy, immune disorders and certain infections. The National Lab Databank contains information on type of test, value, unit, and date of sample collection.

Only laboratory tests encoded with Nomenclature for Property and Unit (NPU) terminology (that is, codes beginning with either “NPU” or “DNK”) is transferred from the National Laboratory Databank to the RLRR. RLRR is updated weekly from the National Laboratory Database.

9.4.2 Sweden

The Swedish National Quality Registry of Assisted Reproduction (Q-IVF) is maintained by a regional health authority, the Region Västra Götaland, and has data since 2007. The Q-IVF covers virtually 100% of IVF treatments in the Swedish population (10.4 million in 2020). It contains information on all IVF and IUI records from all IVF clinics in Sweden (currently, 6 public and 13 private). The information in the register includes a range of patient characteristics, all IVF and insemination treatments and their results (including pregnancy outcome). Approximately 20,000 IVF treatments per year are included in the register. Between 1982 and 2006, data was directly reported to the National Board of Health and Welfare with only clustered data (*i.e.*, without personal identification numbers). The Q-IVF used to be updated on a yearly basis, with data on IVF treatments and treatment outcomes becoming available approximately 2 years after the treatment was performed. Starting from 2020, all participating clinics report data on a regular, more frequent basis (33). Data availability lag time may therefore be reduced. Access to the register data for scientific research is applied from the regional health authority, after a dialogue with the authority. Applications are reviewed by a steering committee of the register.

The Swedish Medical Birth Register (SMBR) is managed by the National Board of Health and Welfare and has data since 1973. The register covers 100% of pregnancies when the child is born (live or stillbirth) after the gestational age of at least 22 weeks since July 2008 (week 28 before July 2008). Thus, no data on elective terminations or spontaneous abortions are available and information on stillbirths is only available from the gestational week 22 onwards (week 28 before July 2008). The register contains information reported by reproductive care, covering the duration of pregnancy and the infant period. The information recorded includes data on the mother (e.g. maternal diagnoses and exposures during pregnancy, including drug use since 1995), the delivery and the neonatal care. The data for the previous calendar year is usually available in December. Access to the register data for scientific research is applied from the National Board of Health and Welfare through a standard application procedure.

The Swedish National Patient Register (SNPR) is managed by the National Board of Health and Welfare. The register covers virtually 100% of the Swedish population (10.4 million in 2020). This register includes all inpatient care since 1987 (ICD-10 from 1998) and specialized outpatient care

from 2001 onwards. It does not cover primary care. The register contains information on, e.g., date of admission and discharge, medical department where the patient was treated, patient demographics, diagnoses and procedures. Diagnoses are coded to ICD-10. Procedures are based on the Nordic Classification of Surgical Procedures (NCSP) and the Nordic Classification of Medical Procedures (NCMP). The register is updated annually with a ca 7-9-month lag-time to data access. The register has increased the reporting frequency from healthcare unites to only monthly, and thus more frequent than annual updates of the register for research purposes and faster data access can be expected in the future. Access to the register data for scientific research is decided by the National Board of Health and Welfare through a standard application procedure and is assessed on a case-by-case basis.

The Swedish Prescribed Drug Register (SPDR) is managed by the National Board of Health and Welfare. It covers virtually 100% of the Swedish population (10.4 million in 2020) and contains data on all drugs prescribed to patients in both primary and secondary care. The register has data since July 2005 and is updated monthly. It contains information on all prescribed medicines dispensed by community pharmacies, e.g., date of prescribing and dispensing, brand name, ATC code, strength and package size of the drug, number of packages and DDDs dispensed, and speciality of the prescriber. The prescribed dose and/or the indication, however, are only available as an optional free-text instruction to the patient. Drugs provided at hospitals or medical clinics (i.e., in the inpatient hospital setting) are not included. The data for a specific month is available with ca 14-day lag-time. Access to the register data for scientific research is applied from the National Board of Health and Welfare through a standard application procedure and is assessed on a case-by-case basis.

The Swedish Cause of Death Register is managed by the National Board of Health and Welfare. This register contains data from 1961 and is updated annually. There is also a historical register of causes of death for the years 1952–1960. The register includes all deaths that have occurred in Sweden. This also includes the death of the person was not registered in Sweden at the time of the death. Key variables include age, sex, ethnicity (country of birth and residency), date of death, cause of death (ICD-9/10 codes have been used since 1988), and intent (in cases of injury or poisoning), medical procedures (if death occurred within 4 months after surgery) and autopsy. Stillbirths are not included in the register. The register is updated annually with a ca 6-9-month lag-time to data access for the cause of death variables, however dates of death are available in the register earlier than that.

9.5 Study Size

The study size calculations are based on publicly available annual reports from the Danish and Swedish IVF registers and the previous feasibility study conducted by PPD (34–37). The available data from the IVF registers included the number of initiated stimulation cycles per year and the number of resulting deliveries. The feasibility study report included information on the number of women who had at least one dispensation of a study drug (follitropin alfa or hMG products) per year. Based on the feasibility study report, there will be thousands of women who have dispensed study drugs at least once during the study period, both in Denmark and Sweden.

The study size calculations considered the primary outcome measures, which are LBR, CLBR and MC-CLBR. As stated in Table 7 in Section 9.7.3.3, the denominator for LBR and CLBR is the

number of initiated stimulation cycles with a given study drug and the denominator for MC-CLBR is women who initiated at least one stimulation cycle with a given study drug. To allow for a sufficient time period to detect live births after the initiation of a stimulation cycle, stimulation cycles with the study drugs initiated on or before 31 December 2019 will be considered in the main analysis of the primary objective.

Total amount of data available for the study and the anticipated proportions between study drug use in the study countries

Using the publicly available yearly summary statistics from the Danish and Swedish IVF registers and the number of women with study drug dispensations from the previous feasibility study report, the total amount of available data in the registers was estimated. For the study period, the total number of all initiated stimulation cycles, and the number of women having at least one stimulation cycle with the study drugs and the proportions of study drug used for stimulation cycles were estimated.

The total number of initiated stimulation cycles over the study periods was estimated by summing up all the yearly reported initiated stimulation cycles in the study periods according to the annual reports from the Danish and Swedish IVF registers. For the study period starting from 2010 (HP-hMG comparison), the estimated total number of initiated stimulation cycles (with any drug) is ~100,000 in Denmark and ~115,000 in Sweden. For the study period starting from 2014 (rhFSH-alfa biosimilars comparison), the estimated total number of initiated stimulation cycles (with any drug) is ~58,000 in Denmark and ~65,000 in Sweden.

The proportion of study drug use over the stimulation cycles was estimated by calculating the proportion of women with at least one study drug dispensation from the total number of started stimulation cycles reported yearly. An assumption was made that one woman initiated one stimulation cycle with one study drug per year. In Denmark, the estimated proportion of study drug use was ~53% (study period of 2010-2019) and ~54% (2014-2019) for rhFSH-alfa reference product; ~33% (2010-2019) for HP-hMG; and ~10% (2014-2019) for rhFSH-alfa biosimilar products. In Sweden, the estimated proportion of study drug use was ~36% (2010-2019) and ~35% (2014-2019) for rhFSH-alfa reference product; ~35% (2010-2019) for HP-hMG; and ~28% (2014-2019) for rhFSH-alfa biosimilar products.

Anticipated amount of data available for the study in the study countries

Using the yearly estimated total number of initiated stimulation cycles together with the proportion of study drug use in the study countries, the anticipated number of initiated stimulation cycles available for the study was estimated.

The total number of women who had used each study drug at least once in a stimulation cycle was estimated by summing up the yearly number of women who had dispensed the study drug at least once. It was assumed that the difference in the yearly numbers of women dispensing the study drug and the number of women with a live-birth delivery corresponded to the number of women who continued using the same study drug the next year. The proportion of women delivering per stimulation cycle was estimated per year by calculating the ratio between the number of deliveries and the number of initiated fresh cycles.

The summary of the total anticipated number of initiated stimulation cycles and the number of women per study drug are summarized in Table 2.

Table 2. Anticipated approximate total number of initiated stimulation cycles and women using drug of interest available per study country and study cohorts[+]

	Study period starting 2010		**Study period starting 2014**	
Country	**rhFSH-alfa reference product**	**HP-hMG products**	**rhFSH-alfa reference product**	**rhFSH-alfa biosimilar products**
Denmark	53 000 cycles	33 000 cycles	31 000 cycles	5 000 cycles
	19 000 women	12 000 women	14 000 women	3 000 women
Sweden	42 000 cycles	40 000 cycles	22 000 cycles	18 000 cycles
	18 000 women	17 000 women	11 000 women	10 000 women

FSH: follicle-stimulating hormone; hMG: human menopausal gonadotropin; HP: highly purified; rhFSH: recombinant human FSH.

Notes:

[+] Cycles refer to the approximate number of initiated stimulation cycles with drug of interest on or before 31 December 2019. Women refer to the approximate number of women who are using drug of interest on or before 31 December 2019.

In Sweden, the ratio between the rhFSH-alfa reference product and HP-hMG cycles and users was close to 1:1, which indicates that this study is anticipated to capture similar number of cycles and women for rhFSH-alfa reference product and HP-hMG study cohorts. The ratio between the initiated stimulation cycles with rhFSH-alfa reference product and rhFSH-alfa biosimilar product was ca 1.25 indicating that the rhFSH-alfa reference product study cohort will be approximately 25% larger than the rhFSH-alfa biosimilar product study cohort, in respect to initiated stimulation cycles. Regarding the number of women with rhFSH-alfa reference product and rhFSH-alfa biosimilar product, the study cohorts is anticipated to capture similar number of users.

In Denmark, the ratio between the rhFSH-alfa reference product and HP-hMG users was ca 1.6, which indicates that the rhFSH-alfa reference product study cohort will be approximately 60% larger than the HP-hMG study cohort. The ratio in the number of women between rhFSH-alfa reference product and rhFSH-alfa biosimilar product was as high as 5:1, which indicates that the rhFSH-alfa reference product study cohort is anticipated to include approximately 5 times as many women as the rhFSH-alfa biosimilar product study cohort.

Study size needed to show a conservative minimum detectable difference between the study cohorts

As seen from the estimations above, the number of women having a stimulation cycle with a study drug is the limiting factor for the study size. This is because each woman included in the analysis will have at least one stimulation cycle. Thus, the final study size estimations consider number of women and uses the primary outcome measure of MC-CLBR.

Assuming the above described ratios in the number of women between the study cohorts for the given comparisons, the conservative minimal detectable difference between the study groups was investigated to show the minimal detectable differences in primary outcomes that can be detected in this study. The minimal detectable difference was calculated by using the measure of association in this study: the OR and its corresponding 95% CI. The minimal detectable difference is the difference in the outcome rates between the study cohorts that will yield an OR whose lower 95% CI limit is above 1. In other words, the minimum detectable difference shows the minimal difference between the outcome rates that shows statistically significant difference between the study cohorts. For the conservative estimate, MC-CLBR of 50% in the comparison drug study cohort is used. Table 3 shows the sample size needed for the different conservative estimates of minimum detectable differences between the study cohorts.

Table 3. The study cohort sizes (number of women) needed to show a conservative minimum detectable difference in MC-CLBR (in percentage point (% point)) between the study cohorts[+]

		Study period starting 2010		Study period starting 2014	
Country	**Minimum detectable difference**	**rhFSH-alfa reference product cohort (women)**	**HP-hMG products cohort (women)**	**rhFSH-alfa reference product cohort (women)**	**rhFSH-alfa biosimilar products cohort (women)**
Denmark	1.0% point	23 000	14 400	53 000	10 600
	1.5% point	12 000	7 500	27 500	5 500
	2.0% point	6 500	4 100	15 000	3 000
	2.5% point	4 100	2 600	9 500	1 900
	3.0% point	2 800	1 800	6 500	1 300
	4.0% point	1 600	1 000	3 500	700
	5.0% point	1 000	600	2 300	450
Sweden	1.0% point	17 600	17 600	17 600	17 600

		Study period starting 2010		Study period starting 2014	
Country	**Minimum detectable difference**	**rhFSH-alfa reference product cohort (women)**	**HP-hMG products cohort (women)**	**rhFSH-alfa reference product cohort (women)**	**rhFSH-alfa biosimilar products cohort (women)**
	1.5% point	8 000	8 000	8 000	8 000
	2.0% point	4 600	4 600	4 600	4 600
	2.5% point	3 000	3 000	3 000	3 000
	3.0% point	2 100	2 100	2 100	2 100
	4.0% point	1 150	1 150	1 150	1 150
	5.0% point	750	750	750	750

HP-hMG: highly purified human menopausal gonadotropin; rhFSH: recombinant human follicle-stimulating hormone; MC-CLBR: multiple-cycle cumulative live birth rate.

Notes:

[+] Women refer to the approximate number of women who are using drug of interest on or before 31 December 2019

To take into account the effect of the study inclusion and exclusion criteria restricting the study population to a subset of the total population available in the registries, a crude assumption that 75% of the women in the registries are eligible for the study was made. With this crude assumption, the anticipated study size will be able to show any differences of 1.5% point in primary outcomes between the rhFSH-alfa reference product and HP-hMG study cohorts, in both Denmark and Sweden. For the comparison of rhFSH-alfa reference product and biosimilars, the anticipated minimum difference that can be detected in primary outcomes will be 2.5% point in Denmark and 2.0% point in Sweden.

Table 4 shows the minimum detectable difference converted from the percentage point (% point) differences into units of percentage difference between the study cohorts at varying MC-CLBR levels in comparator cohort. Assuming MC-CLBR of 25% or higher in the comparator study cohort, 1.5% point, 2.0% point and 2.5% point differences would correspond to 6.0%, 8.0% and 10.0% or lower percentage differences between the study cohorts. Based on clinical judgment, a difference of 5% between the cohorts is deemed clinically meaningful.

Table 4. The minimum detectable difference converted from percentage point (% point) differences into units of percentage difference between the study cohorts for varying MC-CLBR in the comparator cohort

	MC-CLBR in comparator cohort										
Minimum detectable difference	**15%**	**20%**	**25%**	**30%**	**35%**	**40%**	**45%**	**50%**	**60%**	**70%**	**80%**
1.0% point	6.7%	5.0%	4.0%	3.3%	2.9%	2.5%	2.2%	2.0%	1.7%	1.4%	1.3%
1.5% point	10.0%	7.5%	6.0%	5.0%	4.3%	3.8%	3.3%	3.0%	2.5%	2.1%	1.9%
2.0% point	13.3%	10.0%	8.0%	6.7%	5.7%	5.0%	4.4%	4.0%	3.3%	2.9%	2.5%
2.5% point	16.7%	12.5%	10.0%	8.3%	7.1%	6.3%	5.6%	5.0%	4.2%	3.6%	3.1%
3.0% point	20.0%	15.0%	12.0%	10.0%	8.6%	7.5%	6.7%	6.0%	5.0%	4.3%	3.8%
4.0% point	26.7%	20.0%	16.0%	13.3%	11.4%	10.0%	8.9%	8.0%	6.7%	5.7%	5.0%
5.0% point	33.3%	25.0%	20.0%	16.7%	14.3%	12.5%	11.1%	10.0%	8.3%	7.1%	6.3%

MC-CLBR: multiple-cycle cumulative live birth rate.

[redacted]

9.6 Data Management

All study permit approvals and access to the study data will be applied for by PPD. Application for access to individual-level data for scientific research follows a standard application procedure at the Danish and Swedish authorities. For Denmark, pseudonymized individual-level data are placed at a secure, remote server of the Danish Health Data Authority. All data analyses using the linked register data must be conducted at the server of the authority by PPD staff personally authorized by the authority under the authorisation held by PPD Solutions Denmark A/S. No other PPD staff will have access to the secure server environment. For Sweden, data from each register will be extracted by the respective register holder and delivered to PPD as pseudonymized individual-level data without unique personal identification numbers. The researchers at PPD in Denmark and Sweden will thus have access to data where individuals cannot be directly identified.

Before PPD can access the data, the data holders have collected and managed data according to their own standards. As the data contained with the national registers have been collected independently of the current study, the data collection process could not affect the research question. Once the data can be accessed, PPD will start processing the data. Each country will perform analyses with individual-level data separately as described in Section 9.7. SAS v9.4 (SAS Institute, Inc. Cary, North Carolina) or R v3.6.5 or higher (R Foundation for Statistical Computing, Vienna, Austria) languages will be used in data management for creating the analysis database, and in statistical analysis for creating tabulations and graphics, as well as in all statistical modelling. The source code of data management and data analyses is kept for inspection for five years after completion of the study (latest of final study report or first publication of study results).

According to PPD standard operating procedure, only selected study staff have rights to files and directories that contain individual-level data stored on site. Access to the individual-level study data cannot be given to any third parties; only aggregated results will be presented to the Sponsor or otherwise published. The study data cannot be used for other purposes than described in this protocol. All requests to use the study data for other purposes must be subjected to appropriate data permit processes.

PPD will maintain information on the study individuals securely on site according to up-to-date standard operating procedures. PPD will also maintain appropriate data storage, including periodic backup of files and archiving procedures. PPD will comply with procedures that include checking electronic files, maintaining security and data confidentiality, following analyses plans, and performing quality checks for all programs.

9.7 Data Analysis

Further details on the analyses corresponding to each study objective are provided in a SAP.

All analyses will be performed separately for each country. MA may be used, as appropriate, to provide summary estimates for the entire study population across both countries (see Section 9.7.3.8).

9.7.1 Analysis Sets

To address the study objectives, analysis sets will be constructed based on the study outcomes as defined in Section 9.1.2. To allow for a sufficient time period to assess the outcomes after the initiation of a stimulation cycle, the analysis sets will be restricted to outcome-specific time periods. For the analysis sets, the following outcome-specific time periods are applied: 1 year after the stimulation cycle initiation for LBR, CPR, OPR, MC-CLBR and TTLB; 1 year after the stimulation cycle and linked FET initiation for CLBR, CCPR, COPR, pregnancy loss, multiple pregnancy and treatment-associated costs; 1.5 months after the stimulation cycle initiation for oocyte retrieval and embryos cryopreserved per initiated stimulation cycle, and cycle cancellation; 1 month after the stimulation cycle aspiration for oocyte retrieval and embryos cryopreserved per aspirated stimulation cycle; 4 months after stimulation cycle initiation for OHSS; 3 months after last embryo transfer for implantation. Fresh and frozen embryo transfers will be assessed by using all fresh and frozen embryo transfers done by the end of the study period (31 December 2020).

The analysis sets and their use for the study objectives and outcomes are defined in Table 5. The use of analysis sets will be described in the statistical analysis, and any needed specific modifications or restrictions on the analysis sets needed for the analysis of the study objectives will be further described in the SAP.

Table 5. Description of analysis sets

Analysis set	Description of the analysis set	Study objective(s)	Study outcome measure(s)
Set 1	Stimulation cycles with drug of interest initiated on or before 31 December 2019	Primary objective	LBR
		Secondary objective 1	CPR
			OPR
Set 2	Stimulation cycles with drug of interest and linked FET cycles initiated on or before 31 December 2019	Primary objective	CLBR
		Secondary objective 1	CCPR
			COPR
		Secondary objective 5	Pregnancy loss
		Secondary objective 6	Multiple pregnancy
		Secondary objective 9	Treatment-associated costs
Set 3	Stimulation cycles with drug of interest initiated on or before 15 November 2020+	Secondary objective 2	Oocytes retrieved per initiated stimulation cycle
			Embryos cryopreserved per initiated stimulation cycle
			Utilizable embryos per initiated stimulation cycle
		Secondary objective 7	Cycle cancellation

Analysis set	Description of the analysis set	Study objective(s)	Study outcome measure(s)
Set 4	Stimulation cycles with drug of interest aspirated on or before 30 November 2020+	Secondary objective 2	Oocytes retrieved per aspirated stimulation cycle
			Embryos cryopreserved per aspirated stimulation cycle
			Utilizable embryos per aspirated stimulation cycle
Set 5	Stimulation cycles with drug of interest initiated on or before 31 August 2020+	Secondary objective 8	OHSS
Set 6	Fresh and frozen embryo transfers linked to stimulation cycles with drug of interest occurring on or before 31 December 2020+	Secondary objective 2	Fresh embryos transferred
			Frozen embryos transferred
Set 7	Fresh and frozen embryo transfers linked to stimulation cycles with drug of interest occurring on or before 30 September 2020+	Secondary objective 3	Implantation rate
Set 8	Women who initiated their first-ever stimulation cycle with drug of interest on or before 31 December 2019	Primary objective	MC-CLBR
		Secondary objective 4	TTLB

CCPR: cumulative clinical pregnancy rate; CLBR: cumulative live birth rate; CPR: clinical pregnancy rate; COPR: cumulative ongoing pregnancy rate; FET: frozen embryo transfer; LBR: live birth rate; MC-CLBR: multiple-cycle cumulative live birth rate; OPR: ongoing pregnancy rate; OHSS: ovarian hyperstimulation syndrome; TTLB: time-to-live birth.

Notes:

[+]Should data on initiated cycles from the Swedish Q-IVF be available only through 31 December 2019, this analysis set will be adjusted to the same day and month in 2019.

9.7.2 Derived and Transformed Data

The data sources included for this study are national healthcare registers that are expected to have a comprehensive coverage and high-quality data. Thus, large amounts of missing data is not expected to be an issue in this study. In case of missing data, a missing data category will be added, and the number of observations and patients with missing data will be reported for each measured variable in the study. The handling of missing data will be described in more detail in the SAP.

The IVF registries may record some of study covariates related to patient characteristics (see Table 1, section 9.3.2), such as smoking status or BMI, only for the patient's first stimulation cycle. Thus, the covariate value from the patient's first stimulation cycle will be used also for the patient's subsequent cycles if the covariate values for the subsequent cycles are missing.

Any study variables and outcomes that are only recorded if an event occurs, such as live birth, are not considered missing. Also, any outcomes relying on the presence of a diagnostic code (i.e. ICD-10), such as OHSS, fit this definition.

Any derived and transformed data needed for conducting the statistical analysis will be described in the SAP.

9.7.3 Statistical Methods

For the primary objective, the main data analysis will be conducted in two stages: (i) construction of the IPTW-weighted study cohorts by modelling the rhFSH-alfa reference product vs. HP-hMG stimulation cycle initiation and rhFSH-alfa reference product vs. rhFSH-alfa biosimilar product stimulation cycle initiation, (ii) estimating the effect of rhFSH-alfa reference product on the (cumulative) live birth rates using adjusted (IPTW-weighted) odds ratios (ORs), comparing rhFSH-alfa reference product study cohort to the comparators (HP-hMG study cohort or rhFSH-alfa biosimilars study cohort). The IPTW-weighted study cohorts will also be used for the secondary objectives to estimate the effect of rhFSH-alfa reference product on the secondary outcomes, comparing rhFSH-alfa reference product study cohort with the comparators (HP-hMG study cohort or rhFSH-alfa biosimilars study cohort).

Table 6 below summarizes the data analysis planned for the study and an overview of the planned analysis is given in the subsections below. Further details of the statistical analysis will be provided in the SAP.

Table 6. Overview of planned statistical analysis per outcome measure

Study objective	Outcome	Unit of analysis	Analysis set	Analysis results
Primary objective	LBR	Per initiated stimulation cycle	Set 1	Rate (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4) Stratified analysis (see section 9.7.3.6)
	CLBR	Per initiated stimulation cycle	Set 2	Rate (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4) Stratified analysis (see section 9.7.3.6)
	MC-CLBR	Per woman	Set 8	Rate (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4)
Secondary objective 1	CPR	Per initiated stimulation cycle	Set 1	Rate (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4) Stratified analysis (see section 9.7.3.6)
	OPR	Per initiated stimulation cycle	Set 1	Rate (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4) Stratified analysis (see section 9.7.3.6)
	CCPR	Per initiated stimulation cycle	Set 2	Rate (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4) Stratified analysis (see section 9.7.3.6)
	COPR	Per initiated stimulation cycle	Set 2	Rate (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4) Stratified analysis (see section 9.7.3.6)
Secondary objective 2	Number of oocytes retrieved	Per initiated stimulation cycle	Set 3	Number (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4) Stratified analysis (see section 9.7.3.6)
		Per aspirated stimulation cycle	Set 4	Number (see section 9.7.3.3)

Study objective	Outcome	Unit of analysis	Analysis set	Analysis results
				Adjusted OR (see section 9.7.3.4) Stratified analysis (see section 9.7.3.6)
	Number of embryos transferred	Per fresh embryo transfer	Set 6	Number (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4)
		Per frozen embryo transfer	Set 6	Number (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4)
	Number of embryos cryopreserved	Per initiated stimulation cycle	Set 3	Number (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4)
		Per aspirated stimulation cycle	Set 4	Number (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4)
	Number of utilizable embryos	Per initiated stimulation cycle	Set 3	Number (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4)
		Per aspirated stimulation cycle	Set 4	Number (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4)
Secondary objective 3	Implantation rate	Per embryos transferred per stimulation cycle	Set 7	Rate (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4)
Secondary objective 4	Time-to-live birth	Per woman per up to five stimulation cycles with the same drug	Set 8	Kaplan-Meier plots and estimates (see section 9.7.3.5) Adjusted live birth rate ratios (see section 9.7.3.5)
Secondary objective 5	Rate of pregnancy loss	Per clinical pregnancy following embryo transfer	Set 2	Rate (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4)
Secondary objective 6	Multiple pregnancy rate	Per clinical pregnancy following embryo transfer	Set 2	Rate (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4)

Study objective	Outcome	Unit of analysis	Analysis set	Analysis results
Secondary objective 7	Cycle cancellation rate	Per initiated stimulation cycle	Set 3	Rate (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4)
Secondary objective 8	OHSS rate	Per initiated stimulation cycle	Set 5	Rate (see section 9.7.3.3) Adjusted OR (see section 9.7.3.4)
Secondary objective 9	Treatment-associated costs	Per initiated stimulation cycle	Set 2	Descriptive cost analysis (see section 9.7.3.7)

CCPR: cumulative clinical pregnancy rate; CLBR: cumulative live birth rate; CPR: clinical pregnancy rate; COPR: cumulative ongoing pregnancy rate; LBR: live birth rate; MC-CLBR: multiple-cycle cumulative live birth rate; OPR: ongoing pregnancy rate; OHSS: ovarian hyperstimulation syndrome; OR: odds ratio.

9.7.3.1 Propensity Score Methodology

PS-based methods are frequently used in non-interventional studies to control for confounding when estimating treatment effects (38). In this study, the PS weighting will be implemented using IPTW approach.

PS, typically obtained using logistic regression models, are covariate summary scores that indicate the individual probability of having a treatment initiated with the study drug against the comparator drug, given the observed baseline characteristics. In PS weighting with IPTWs, each subject is weighted by the inverse probability of receiving the treatment they have actually received, *i.e.*, 1/PS in the study drug cohort and 1/(1-PS) in the comparator drug cohort (22). In other words, the PS weighting with IPTW creates a pseudo-population of study drug cohort and the comparator drug cohort that are weighted samples of the original study cohorts but have the same covariate distribution as the overall original study cohorts (22). Thus, using the IPTW-weighted study cohorts will remove confounding from all covariates included in the PS model, and give the estimate for the treatment effect in the overall population (22).

For this study, the IPTW approach will be based on pairwise PS estimation for i) rhFSH-alfa reference product cohort and HP-hMG cohort, and ii) rhFSH-alfa reference product cohort and rhFSH-alfa biosimilars cohort. The PS will be estimated for the stimulation cycle initiation with the study drug (rhFSH-alfa reference product) against the comparator drug (HP-hMG or rhFSH-alfa biosimilar product) by using logistic regression, including covariates as indicated in Table 1 (Section 9.3.2). Based on the estimated pairwise PS, i) IPTW-weighted rhFSH-alfa reference product cohort and IPTW-weighted HP-hMG cohort, as well as ii) IPTW-weighted rhFSH-alfa reference product cohort and IPTW-weighted rhFSH-alfa biosimilars cohort will be created.

Diagnostic measures will be used for assessing the success of the PS model to account for confounding. In addition, the distribution of IPTWs will be checked and in case very large weights are observed, the stabilized, trimmed and truncated weights will be considered to avoid increasing the variability of the treatment effect estimates due to very large weights (38).

PS weighting and evaluation of its success will be done separately per each analysis set (see Section 9.7.1) as covariate balance is not automatically the same in the analysis sets that include data from slightly varying time periods and thus have may have slightly varying set of data included. Also, for stratified analyses, PS weighting and evaluation of its success will be re-done as covariate balance is not automatically maintained in the sub-groups that are created by the stratification.

In addition, other PS implementation methods, such as matching or stratification, may be considered in case the received data does not support for using IPTW approach, for any reason.

9.7.3.2 Descriptive Analysis

Descriptive analysis will be conducted to describe the baseline characteristics of the study cohorts. For descriptive analysis of continuous variables, the number of observations, number of missing values, mean, standard deviation, median, lower and upper quartiles, as well as 5th and 95th percentiles will be presented. For categorical variables, numbers and percentages of patients for each of the categories and numbers and percentages of missing values will be presented in descriptive analysis.

9.7.3.3 Rates and Numbers

The results of the study outcomes in the study cohorts will be estimated by calculating the outcome measures according to their definitions (see Table 7) in both original and IPTW-weighted study cohorts.

For the analyses of LBR, a woman may contribute up to one delivery resulting in live birth. For the analyses of CLBR, a woman may contribute several deliveries resulting in live birth per each stimulation cycle. The same principle applies to the analyses of clinical and ongoing pregnancies, *i.e.*, a woman may contribute several clinical and ongoing pregnancies for the analyses of CCPR and COPR, respectively.

For the analysis of MC-CLBR, where the woman is the unit of analysis, only the first occurrence of a live birth per woman within the first five stimulation cycles using the same study drug will be considered. The MC-CLBR analysis will be performed by cycle rank, *i.e.*, analyzed as the cumulative probability of a live birth up to and including the first, second, third, fourth and fifth stimulation cycle, respectively. The MC-CLBR analysis will be restricted to women who initiated their first-ever stimulation cycle during the study period.

In the assessment of MC-CLBR, it is assumed that women who did not return during the study period for another stimulation cycle after a failed attempt (*i.e.*, who did not have a live birth), did not have any live birth. Thus, for the main analysis, the MC-CLBRs will be calculated as conservative estimates (20). Optimal estimates will be provided in a sensitivity analysis restricted to the primary objective (see Analysis 2 Section 9.7.3.9).

Rates and numbers will be expressed as the rate or number per 100 units of observations with 95% CI.

Table 7. Definitions of outcome measures

Study objective	Outcome measure	Numerator	Denominator
Primary objective	LBR	Number of deliveries with at least one live birth after fresh embryo transfer	Initiated cycles with drug of interest
	CLBR	Number of deliveries with at least one live birth after one initiated stimulation cycle, including all fresh or frozen embryo transfers linked to that cycle	Initiated cycles with drug of interest
	MC-CLBR	Number of deliveries with at least one live birth after each initiated stimulation cycle in up to five cycles, including all fresh or frozen embryo transfers linked to each of those stimulation cycles	Women who initiated 1, 2, 3, 4 or 5 cycles with drug of interest
Secondary objective 1	CPR	Number of clinical pregnancies after fresh embryo transfer	Initiated cycles with drug of interest
	CCPR	Number of clinical pregnancies after one initiated stimulation cycle, including all fresh or frozen embryo transfers linked to that cycle	Initiated cycles with drug of interest
	OPR	Number of ongoing pregnancies after fresh embryo transfer	Initiated cycles with drug of interest
	COPR	Number of ongoing pregnancies after one initiated stimulation cycle, including all fresh or frozen embryo transfers linked to that cycle	Initiated cycles with drug of interest
Secondary objective 2	Number of oocytes retrieved	Number of oocytes retrieved	Initiated cycles with drug of interest
		Number of oocytes retrieved	Aspirated cycles with drug of interest

Study objective	Outcome measure	Numerator	Denominator
	Number of embryos transferred	Number of fresh embryos transferred	Fresh embryo transfers linked to cycles with drug of interest
		Number of frozen embryos transferred	Frozen embryo transfers linked to cycles with drug of interest
	Number of embryos cryopreserved	Number of embryos cryopreserved	Initiated cycles with drug of interest
		Number of embryos cryopreserved	Aspirated cycles with drug of interest
	Number of utilizable embryos	Number of utilizable embryos	Initiated cycles with drug of interest
		Number of utilizable embryos	Aspirated cycles with drug of interest
Secondary objective 3	Implantation rate	Number of intrauterine gestational sacs after one fresh or frozen embryo transfer	Number of embryos transferred in one cycle
Secondary objective 4	TTLB	Time to first delivery with at least one live birth in up to five stimulation cycles with the same study drug, including all fresh or frozen embryo transfers linked to each of those stimulation cycles	N/A; conducted as time-to-event analysis; unit of analysis is women who initiated stimulation cycles with drug of interest
Secondary objective 5	Rate of pregnancy loss	Number of induced and spontaneous abortions	Clinical pregnancies with drug of interest following fresh or frozen embryo transfer
Secondary objective 6	Multiple pregnancy rate	Number of multiple pregnancies	Clinical pregnancies with drug of interest following fresh or

Study objective	Outcome measure	Numerator	Denominator
			frozen embryo transfer
Secondary objective 7	Cycle cancellation rate	Number of canceled cycles	Initiated cycles with drug of interest, including any linked initiated FET cycles
Secondary objective 8	OHSS rate	Number of OHSS cases	Initiated cycles with drug of interest

CCPR: cumulative clinical pregnancy rate; CLBR: cumulative live birth rate; CPR: clinical pregnancy rate; COPR: cumulative ongoing pregnancy rate; FET: frozen embryo transfer; LBR: live birth rate; MC-CLBR: multiple-cycle cumulative live birth rate; OPR: ongoing pregnancy rate; OHSS: ovarian hyperstimulation syndrome; TTLB: time-to-live birth.

9.7.3.4 Adjusted Odds Ratios

For the comparison of the study cohorts, the ORs (with 95% CI), adjusted for confounding factors, will be estimated by taking into account the IPTWs and using statistical modelling, such as regression analysis or general estimating equations. To estimate the adjusted ORs, the statistical model will be weighted with IPTWs and will include as covariates any variables that were used in the PS but are still unbalanced between the study cohorts after weighting. Statistical models may also be used to additionally adjust the ORs for any other potential covariates that are assessed after the initiation of the IVF/ICSI stimulation cycle and hypothesized to be associated with the given study outcomes (see Table 1 in Section 9.3.2).

For the statistical models, the following study outcomes will be considered as binary outcomes analyzed per fresh embryo transfer: live birth (*i.e.*, a primary objective outcome), clinical pregnancy and ongoing pregnancy (*i.e.*, a subset of secondary objective 1 outcomes). For the statistical models, the following study outcomes will be considered as binary outcomes analyzed per woman: MC-CLBR (*i.e.*, a primary objective outcome) will be analyzed as a binary outcome per woman in statistical models.

For the statistical models, the following study outcomes will be considered as binary outcomes analyzed per clinical pregnancies following fresh or frozen embryo transfer: pregnancy loss (secondary objective 5) and multiple pregnancy (secondary objective 6). Cycle cancellation (secondary objective 7) and OHSS occurrence (secondary objective 8) will be analyzed as binary outcomes per initiated stimulation cycle. Implantation (secondary objective 3) will be analyzed as an ordinal categorical outcome per fresh or frozen embryo transfer.

For the statistical models, the following study outcomes are considered as counts analyzed per initiated stimulation cycle: cumulative live birth (*i.e.*, a primary objective outcome), cumulative clinical pregnancy and cumulative ongoing clinical pregnancy (*i.e.*, a subset of secondary objective 1 outcomes). Also the following study outcomes are considered as counts in statistical models and

will be analyzed both per initiated stimulation cycle and per aspirated stimulation cycle: number of oocytes retrieved and number of utilizable embryos (*i.e.*, a subset of secondary objective 2 outcomes). Number of embryos transferred (*i.e.*, a secondary objective 2 outcome) will also be considered as a count outcome in statistical models and will be analyzed per fresh and frozen embryo transfers. For modelling the count data appropriately, the distribution of count data will be checked by counting the mean and the variance of the data and the possibility of excess zeros in the count data will be evaluated visually by plotting the distribution of the data.

9.7.3.5 Time-to-Live Birth Analysis

TTLB (secondary objective 4) will be analyzed restricting to women who initiated their first-ever stimulation cycle during the study period. As per study design, only the first five stimulation cycles (including all cycles of fresh or frozen embryo transfers) with the same study drug will be considered. The stimulation cycle rank number will be employed as the unit of time. Women will be followed until the first delivery with at least one live birth linked to treatment with the drug of interest, the sixth stimulation cycle, switch to another drug used for COS, diagnosis of breast, ovarian or uterine cancer, death, emigration (Denmark only) or end of study, whichever occurs first.

To estimate the TTLB in the original and IPTW-weighted study cohorts, survival analysis will be performed with the Kaplan-Meier method to estimate the time to the first live birth, expressed as the rank number of the stimulation cycle with the given study drug. The resulting Kaplan-Meier plots together with the estimated mean (with 95% CI) and median (with 95% CI) time-to-events will be presented.

For the comparison of the study cohorts, the time-to-live birth rate ratios adjusted for confounding factors, will be estimated by taking into account the IPTWs and using Cox proportional hazards regression model. The Cox model will be weighted with the IPTWs, and any variables included in the PS that are still unbalanced after weighting will be included as covariates for the Cox model. Additionally, the Cox model may be used to adjust the rate ratios for any other potential covariates that are assessed after the start of the stimulation cycle and hypothesized to be associated with the given study outcomes (see Table 1 in Section 9.3.2).

9.7.3.6 Stratified Analysis

For the stratified analysis, the study population will be stratified into groups based on their baseline characteristics of age (<35 years old vs. ≥ 35 years old) and use of GnRH agonist or antagonist protocol. The stratified analysis will be performed for the following outcome measures: LBR, CLBR, MC-CLBR, CPR, CCPR, OPR, COPR and OHSS, and for the number of oocytes retrieved and the number of utilizable embryos per initiated stimulation cycle and per oocyte retrieval cycle. Rates and adjusted odds ratios will be estimated as described for each outcome in Sections 9.7.3.3 and 9.7.3.4.

9.7.3.7 Descriptive Cost Analysis

The treatment-associated costs will be reported descriptively for the initiated IVF/ICSI stimulation cycles for each drug of interest, and overall, irrespective of drug. The costs will be presented per

cycle and per woman. Costs will be reported by category: treatment costs (drugs and other treatments), costs of miscarriage/birth, and adverse event costs. Costs will also be reported in total (sum of cost categories). The descriptive analysis will present the number of observations, mean, standard deviation, median, range, lower and upper quartiles. Missing observations will be reported. Costs will be reported in local currency and in Euros, using relevant exchange rates. Further, costs will be adjusted to the same calendar year using relevant consumer price indexes.

9.7.3.8 Meta-analysis

In addition to the results reported separately for each country, results will also be combined for a more holistic view and overall increased study precision. Because individual-level data currently cannot be shared across the borders of Denmark and Sweden, an MA approach based on aggregate data will be considered. The MA will be performed for the primary study objective only.

Prior to conducting MA, heterogeneity of the country-specific results will be assessed using:

- The estimated total heterogeneity.
- The Chi-squared test (significance level 0.1).
- I^2 statistic (0% to 40%: might not be important; 30% to 60%: may represent moderate heterogeneity; 50% to 90%: may represent substantial heterogeneity; 75% to 100%: considerable heterogeneity).

If high levels of heterogeneity are observed (I^2 >=50% or P<0.1), the characteristics of the data source-specific results will be reviewed, and the possible sources of heterogeneity discussed. Results of the MA will be derived using both fixed and random-effects models, and the main model will be determined by the heterogeneity assessment. MA and country-specific results will be presented in forest plots. Further details will be presented in the SAP.

9.7.3.9 Sensitivity Analyses

Sensitivity analyses will be performed as outlined below for the primary study objective and, in one analysis (no. 3), for secondary objective 4 (Table 8).

Table 8. Overview of sensitivity analyses

No.	Topic	Main analysis		Sensitivity analysis	
		Method	**Justification**	**Alteration to method**	**Justification**
1	Time window for establishing stimulation cycle exposure to study drug	30 days before to 35 days after stimulation cycle index date	The IVF registers do not capture drug data comprehensively. To establish exposure of a stimulation cycle to a study drug, and one study drug only, data on dispensed prescriptions for gonadotropins are ascertained from the prescribed drug registers. A time window has been defined to capture all relevant dispensed prescriptions for gonadotropins in each cycle.	15 days before to 15 days after stimulation cycle index date	The time window for exposure establishment used in the main analysis has only been used in one prior study (27) to our knowledge. The effect of choosing a narrower time window will be explored for the primary objective and each associated outcome measure (LBR, CLBR, MC-CLBR). Otherwise the analysis will be repeated as described in 9.7.3.3 and 9.7.3.4.
2	Approach for estimation of MC-CLBR	The main analysis uses a conservative approach to estimating the MC-CLBR. It assumes none of the women who did not return for a subsequent cycle would have had a live birth, had they returned (20).	Discontinuation of IVF/ICSI treatment, *i.e.*, not returning for treatment after a failed attempt (live birth was not achieved) needs to be considered in the analysis of cumulative live birth rates assessed over multiple stimulation cycles.	Optimal approach to estimating the MC-CLBR. It is based on the observed data and assumes women who did not return for a subsequent cycle would have had the same live birth rates as those who did return (20).	The true MC-CLBR is thought to lie between the conservative and optimal estimates (39). Calculating both estimates will provide the range in which the actual value would be expected to fall. The analysis of MC-CLBR will otherwise be repeated as described in 9.7.3.3 and 9.7.3.4.
3	Minimum duration of follow-up for assessment of MC-CLBR and TTLB	The main analyses of the MC-CLBR and TTLB each uses a minimum of 1 year of follow-up and includes women who had initiated their first-ever stimulation	1 year is considered sufficient to achieve at least one live birth and a reasonable number of complete cycles for reliable estimates of the MC-CLBR in view of previous studies considering up to 9 (20) and 8 (21) stimulation cycles for the estimation of cumulative live birth rates with at least 1.5 and 2 years of follow-up, respectively.	Minimum follow-up period extended to 2 years, thus restricting the analysis to women who had initiated their first-ever stimulation cycle on or before 31 December 2018.	The effect of extending the minimum follow-up period on the conservative and optimal estimates of the MC-CLBR and on TTLB will be explored. The analysis of each outcome will otherwise be repeated as described in 9.7.3.3 and 9.7.3.4.

No.	Topic	Main analysis		Sensitivity analysis	
		Method	**Justification**	**Alteration to method**	**Justification**
		cycle on or before 31 December 2019.			
4	CLBR	The main analysis does not consider the possible impact of oocyte yield on CLBR.	Oocyte yield is not considered in the stratified analysis.	CLBR in the first, second and third stimulation cycle will be estimated by the number of oocytes retrieved in the first cycle (cycle rank expressed as observed in the study period).	Oocyte yield is known to be positively associated with CLBR (20,40,41). The effect of oocyte yield in the first cycle will be explored for the CLBR. Otherwise the analysis will be repeated as described in 9.7.3.3 and 9.7.3.4, although rates will only be reported for the IPTW-weighted cohorts.
5	Linkage between stimulation and FET cycles for the estimation of CLBR and MC-CLBR (Denmark only)	FET cycles preceded by multiple stimulation cycles, of which more than one resulted in cryopreserved embryos, are handled as follows: - If all preceding stimulation cycles used the same gonadotropin, it is assumed that the FET cycle is linked to the most recent stimulation cycle - If the preceding stimulation cycles used different gonadotropins, the FET cycle is	The DIVF does not record a link between an FET cycle and the stimulation cycle it originated from. Because clinics allowed for cryopreserved embryos to remain in the freezer while the woman started a new stimulation cycle, the origin of an FET cycle cannot be specified if it was preceded by multiple stimulation cycles which resulted in cryopreserved embryos.	Two alternative approaches will be taken for FET cycles preceded by multiple stimulation cycles which resulted in cryopreserved embryos: a) FET cycles preceded by stimulation cycles which used the same gonadotropin will be excluded b) The study period will be restricted to 2016-2020, and it will be assumed that an FET cycle is linked to the most recent stimulation cycle.	The effect of altering the FET cycle linkage method will be explored for the CLBR and MC-CLBR in order not to introduce bias by systematically including in the analysis repeat stimulation cycles in which the same gonadotropin was used (approach a). To explore the practice in more recent years in which clinics are less prone to allow banking of frozen embryos, the CLBR and MC-CLBR will be estimated using data from 2016-2020 (approach b). Otherwise the analysis for each approach will be repeated as described in 9.7.3.3 and 9.7.3.4. As described in Section 9.2.2, the assumptions made for the main and sensitivity analyses are subject to modification if deemed inadequate.

No.	Topic	Main analysis		Sensitivity analysis	
		Method	**Justification**	**Alteration to method**	**Justification**
		excluded from all analyses (If an FET cycle was preceded by one stimulation cycle only or multiple stimulation cycles of which only one resulted in cryopreserved embryos, that FET cycle can be reliably linked to the appropriate stimulation cycle.)			
6	Adjustment for laboratory test results (Denmark only)	Adjusted analyses will not consider adjustment for relevant laboratory test results.	Completeness of data from the RLRR is uncertain. It is expected that laboratory test results will be available only for a subset of the Danish study population.	An adjusted analysis will be performed among the subset of Danish stimulation cycles or women, depending on the outcome measure, with at least one record of a relevant laboratory test result (see Table 1) available before the stimulation cycle index date (per cycle analyses: LBR, CLBR) or the index date (per woman analysis: MC-CLBR). This adjusted analysis will include relevant	The effect of controlling for relevant predictors of live birth, such as anti-Müllerian hormone levels, will be explored in this analysis. The adjusted analysis of the primary objective and each associated outcome measure (LBR, CLBR, MC-CLBR) will otherwise be repeated as described in 9.7.3.4.

No.	Topic	Main analysis		Sensitivity analysis	
		Method	Justification	Alteration to method	Justification
				laboratory test results as independent variables in the appropriate regression models.	

ICSI: intracytoplasmic sperm injection; IPTW: inverse probability of treatment weighting; IVF: in vitro fertilization; MC-CLBR: multiple-cycle cumulative live birth rate; RLRR: Danish Register of Laboratory Results.

9.7.4 Sequence of Analyses

The study analysis and results (statistical tables only) conducted using the Danish data are planned to be delivered as interim analysis. The Danish data is expected to be received and analyzed prior to the Swedish data. Therefore, in order to deliver the study results to the Sponsor as soon as result for Denmark is available, the interim analysis with Danish data only was planned. However, the data application processing time for Denmark has increased significantly due to Coronavirus Disease 2019 (COVID-19). With the estimate made available by end of 2020, the data application period will be similar for Sweden and Denmark. In this case, there might not be a need for an interim report. PPD is conducting an ongoing monitoring of update from Danish authorities regarding this update.

9.8 Quality Control

The study protocol has been written by following the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Code of Conduct (42) and the Guideline for Good Pharmacoepidemiology Practices (GPP) by the International Society for pharmacoepidemiology (43). The study protocol and as well as results will be published in the EU electronic register of post-authorization studies (EU PAS register) maintained by the EMA.

At the study level, all aspects of the study from protocol development to the reporting of the results are conducted within the work-frame of PPD Quality Management System (QMS) and in accordance to the following policies and procedures:

- POL_QA_001 “Quality Management System” policy
- POL_QC_001 “Quality Control Strategy” policy
- SOP_QC_002 “Quality Control of Project Deliverables”

According to the policies and procedures above, a Quality Control plan for the study will be developed and executed, which will include quality control on study methodology, statistical analysis plan, programming, data management and analysis, and study report including study results and conclusions. Furthermore:

- The study Quality Control plan will establish ownership for the execution of the individual Quality Control steps. The principle of the independence of Quality Control applies.
- PPD project management will ensure that individuals responsible for the execution of specific Quality Control steps will have knowledge, capability and experience which are adequate for the task.
- The result of the execution of the individual steps of the Quality Control plan will be documented, and include the required corrective actions, if any.
- The execution of any required corrective action will be documented.

- The executed Quality Control plan will be subjected to a final review and approval for sufficiency and completeness by the PPD project management team.

- The project management will also ensure that PPD employees assigned to the project are trained on protocol and project-specific procedures, as per PPD procedure RWI_WI_PM0035 “Real-World Project Specific Training and Staff Transition”.

The study will be conducted as specified in this protocol. All revisions to the protocol must be approved by the principal investigator and the Sponsor. All changes to the protocol shall be properly documented as protocol amendments and when necessary such protocol amendments are delivered to source data register holders.

Quality control will also be performed on the retrieved register data, including controlling the inclusion of compulsory variables, such as study identification number (SID), hospitalization, and the main diagnosis, in the delivered register data. If data are missing or incorrect, the dataset is sent back to the register holder for correction.

All programs for data management and data analyses will be written by the study statistician(s). A quality control check of these programs will be carried out by a statistician other than the one who writes the program. All processes from data management leading to dissemination of study results will undergo quality control checks for programs, result tables, and written text. A detailed audit trail of all documents (programs, result tables, reports) along with quality control processes will be maintained.

9.9 Limitations of the Research Methods

9.9.1 Selection Bias

Owing to the use of secondary data originally collected for purposes unrelated to this study, biased results arising from how women are selected into the study is considered a minor issue. Selection will not depend on both exposure status and risk factors for the outcomes, which can otherwise lead to biased exposure-outcome associations. Moreover, all registers used in the study have nationwide coverage and employ mandatory reporting for publicly- and privately-operated units. The Swedish IVF register is an exception as being a quality register, to which reporting is not mandatory; however, close to 100% of all IVF clinics in Sweden report data to the register (33).

9.9.2 Information Bias

Information on the choice of gonadotropin drug for the stimulation cycles will not be comprehensively available from the IVF registers in the two countries. Therefore, this information will be ascertained based on linked prescriptions from the prescribed drug registers. In the main analyses, prescriptions filled during the time interval from 30 days before to 35 days after stimulation start will be used to classify patients according to study drug. If this interval would be too long, drug prescriptions could be incorrectly attributed to stimulation cycles observed in the IVF register, leading to exposure misclassification. Correspondingly, if the interval would be too short, the actual treatment may not be identified, leading to under-ascertainment of exposure. In a

sensitivity analysis of the primary objective, the time interval to identify treatment will therefore be varied, to assess the impact of different definitions on the results (see Analysis 1 in Table 8).

In the DIVF, no direct link exists between an FET cycle and the stimulation cycle it originated from. Different practices have been used by Danish IVF clinics over the years, where in the early part of the study period (until approximately 2015), clinics would allow women to bank frozen embryos while starting a new stimulation cycle (PPD , personal communication by email, PPD). For the estimation of the CLBR and the MC-CLBR, it is important that FET cycles are linked to the correct stimulation cycle. In sensitivity analyses (see Analysis 5 in Table 8), the methods for linking FET cycles to the appropriate stimulation cycle will be varied.

The exposure classification will be based on filled drug prescriptions. However, a filled prescription does not guarantee that the woman actually injected the drug or otherwise adhered to the injection schedule or prescribed dose. Non-use of dispensed drugs and non-adherence would lead to exposure misclassification. Assuming that misclassification was non-differential between the study cohorts, this would bias the results toward the null (no effect). If misclassification was differential, this could bias the results in any direction. However, considering that drug treatment is part of a wider treatment protocol and that women initiating IVF treatment are likely generally well-motivated and well-counselled, the potential for exposure misclassification should be limited.

In Denmark, a small volume of sales of the study drugs were made to hospitals during the period 2010-2019, which means that some cycles may not be included in the analysis. However, since the percentage of in-hospital sales constituted less than 2.5% and less than 1% of total sales for HP-hMG and follitropin alfa (*i.e.*, rhFSH-alfa reference product and rhFSH-alfa biosimilar products combined), respectively, it is expected to have a minimal impact on the study (44).

In the Nordics, the validity and completeness of the medical birth registers for use in epidemiological research is considered to be high (45), and it is expected that obtaining information on the primary outcome, live birth, will be nearly 100% complete. In the SMBR, for example, it is estimated that 97-99% of all newborn babies in Sweden are reported to the register, and the number of children reported to the SMBR are validated against population registration data (46). Further, the Q-IVF contains over 240,000 registered treatments to date from all public and private IVF-treating clinics in Sweden and is considered one of the most complete, national IVF registers in the world. In the recently published 2020 Q-IVF Annual Report, which reported data from year 2018, there were only 13 cases of initiated fresh IVF cycles with missing data on treatment or pregnancy outcomes; accordingly, the proportion of treatments lost to follow up amounted to only 0.14% of the total 11,061 annual fresh IVF cycles (33). In 2017, a similarly very low proportion (0.13%) lost to follow up was reported in the 2019 Q-IVF Annual Report (47). Researchers suggest that loss of follow up in these few cases may be because patients came from/have moved abroad. However, as more than 99% of all initiated IVF treatments contain sufficient follow up to assess study outcomes, the few cases lost to follow up is expected to have a minimal impact in the study. The Q-IVF has been used widely in research previously (41,48–51).

Similar to Sweden, in Denmark, the DMBR has a high completeness, particularly for the primary outcome of live birth, since all newborn children are registered in the DMBR (52–54). In the early period when the DIVF was first established (1994-2005) and included only paper-based reports,

the Danish National Board of Health suspected that were to some degree an under ascertainment of the true number of treatments initiated due to incomplete capture, noting that the use of electronic capture in future updates would improve quality (55,56). However, as only electronic reports are included during the period under study (2010-2020), previously identified limitations related to underreporting would not be an issue in this study. All treatment cycles and measurements on pregnancies from private and public IVF clinics are required to be reported, and the DIVF during the study period is based on electronic capture and complete and validated data (52). Similar to the Q-IVF, the DIVF has been used extensively in research (27,40,54,57,58).

As in any study assessing cumulative live birth rates of IVF treatment, uncertainty remains around end outcomes in those who discontinued treatment, had they continued. This uncertainty will be addressed by the estimation of conservative estimates and, in a sensitivity analysis, optimal estimates for the primary objective MC-CLBR outcome, based on different assumptions of live birth rates in women who discontinued treatment (see Analysis 2 in Table 8). Similarly, the impact on the conservative and optimal MC-CLBR estimates of extending the minimum follow-up to 2 years will also be explored (see Analysis 3 in Table 8).

Further, prescribing patterns of the study drugs are likely to vary to some extent across the study period. If one drug was favored over the other in the early part of the study period, the probability of observing all live births in up to five stimulation cycles using that drug would be higher than for the other one (due to availability of data with longer follow-up). Based on the feasibility assessment of the use of the study drugs in Denmark and Sweden, it is known that the use of rhFSH-alfa biosimilar products has increased gradually, with a relatively larger proportion observed during the latter part of the study period, compared with rhFSH-alfa reference product (37). Hence, it cannot be ruled out that this may lead to a downward bias for the cumulative live birth rate outcomes following COS with rhFSH-alfa biosimilar products.

Prior to October 2016, induced abortions were not recorded in the SNPR, meaning ascertainment of this component of pregnancy loss will rely solely on the Q-IVF prior to this date. Furthermore, abortion statistics based on anonymized data show that ~70% of all induced abortions between 2018 and the second quarter of 2020 were medical abortions completed at home (59). This suggests the data on induced abortions captured in the SNPR in 2016-2020 is likely incomplete.

9.9.3 Confounding by Baseline Characteristics

Confounding by baseline characteristics is a key challenge in non-interventional studies of drug treatment effects. Although the study drugs are equally recommended in IVF treatment guidelines and used interchangeably, differences in baseline characteristics, e.g., in terms of cause of infertility or previous IVF treatment, between the study cohorts cannot be ruled out. To minimize the potential for confounding, IPTW methods accounting for a broad range of characteristics, including history of IVF treatment, fertility-related health characteristics, and other factors potentially associated with the study outcomes will be applied.

9.9.4 Unmeasured Confounding

Although a broad range of characteristics will be collected for this study, some relevant predictors of treatment outcomes, such as the antral follicle count and anti-Müllerian hormone levels (60,61),

will not be available from the data collected in Sweden and likely only available for a subset of the Danish study population. Unmeasured confounding cannot be ruled out. However, a sensitivity analysis performed for Denmark will explore the effect of controlling for anti-Müllerian hormone levels and other relevant outcome predictors available from the RLRR (see Analysis 4 in Table 8).

9.9.5 Study Size and Time Trends

The study size depends on the inclusion period and the incidence of fertility treatments in both countries of study. Differences in standard of care, treatment patterns and success rates are expected to some extent over time. Legal regulations on eligibility for IVF treatment (e.g., donor gametes, lesbian couples, single women) and rules for storage of frozen oocytes and embryos have also changed during the study period. However, with an overall study period spanning 11 years from 2010-2020, any changes in standard of care are expected to have limited impact on the consistency of results across the study period.

9.9.6 Generalizability

The study results are expected to be representative of similar populations and other countries with a universal, one-payer health care system with similar offerings of fertility treatments, and a support system for women/families with fertility issues.

9.10 Other Aspects

9.10.1 Independent Ethics Committee or Institutional Review Board

A procedure for obtaining informed consent is not applicable for Denmark since informed consent by Danish law is not required for register-based studies using secondary data.

According to Danish law, using de-identified data from Danish health care authorities, the study is categorized as exempt from an Institutional Review Board review and from an Ethics Committee approval. The results of studies – containing subjects' unique pseudonymized identifying number, relevant medical records, and information on dates of birth/death and emigration/immigration – will be recorded and stored at the secure server of the Danish authorities. Patient-level data may not and cannot be transferred to, and used in, other countries. Only aggregated and 100% anonymized data may leave the secure, remote server at the Danish authorities for transferal to the Sponsor.

Prior to commencement of the study in Sweden, the protocol will be submitted to the Swedish Ethical Review Authority for approval. The written approval of the Authority will be filed by the investigator and a copy will be sent to the Sponsor. The Ethical Review Authority will be asked to provide documentation of the date of the meeting at which the approval was given, and of the members and voting members present at the meeting. Written evidence of approval that clearly identifies the study and the protocol version should be provided. Where possible, copies of the meeting minutes should be obtained.

Amendments to the protocol will also be submitted to the Swedish Ethical Review Authority before implementation in case of substantial changes.

9.10.2 Quality Assurance

In compliance with regulatory requirements, the Sponsor, a third party on behalf of the Sponsor, or regulatory agencies may conduct quality assurance audits/inspections at any time during or following a study. The investigator must agree to allow auditors/inspectors direct access to all study-related documents, including source documents, and must agree to allocate his or her time and the time of his or her study staff to the auditors/inspectors in order to discuss findings and issues.

The protocol, each step of the data capture procedure, and the handling of the data, as well as the study report, may be subject to independent Clinical Quality Assurance. Audits may be conducted at any time during or after the study to ensure the validity and integrity of the study data.

9.10.3 Archiving

For Denmark, all source data and analytical data sets that are not 100% anonymous cannot leave the remote, secure server of the Danish authorities. All original and enriched data sets used for the analyses must be stored at the site of the Danish authorities for the longest possible time permitted by the applicable regulations. In any case, PPD should ensure that no destruction of data sets used for analyses is performed without the written approval of the Sponsor.

For Sweden, all documents and material pertaining to the project will be securely stored on site for the whole duration of the study and up to five years after study completion (final study report or first publication of study results, whichever comes later), unless otherwise agreed with the Sponsor or required by local regulation.

10 Protection of Human Subjects

10.1 Subject Identification and Privacy

This study involves analysis of secondary data and all data are pseudonymized to protect the privacy of subjects and providers.

For further protection of subjects' privacy, minimum and maximum values will not be reported. Cells in tables must include a minimum of 5 observations and summary statistics can only be calculated if a minimum of 5 observations are used.

11 Management and Reporting of Adverse Events

According to EU Good Pharmacovigilance Practices (GVP) module VI - Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2), section VI.C.1.2.1.2, 28 July 2017 (62), non-interventional studies based on secondary use of data do not require reporting of suspected adverse reactions in the form of Individual Case Safety Reports. Safety outcomes will be summarized in the final study report as reported in Section 9.7, as well as in the interim study report presenting only the results from Denmark.

12 Plans for Disseminating and Communicating Study Results

12.1 Study Report

PPD will summarize the results of the data from Denmark in an interim study report following the same structure as the final study report but presenting only the results from Denmark. The completed final study, including results from Denmark and Sweden, will be summarized in a final report that accurately and completely presents the study objectives, methods, results, limitations of the study, and interpretation of the findings.

12.2 Publication

Based on the study report, the principal investigator and co-investigators (together referred to as "investigators"; who have contributed to the design of the study protocol) will prepare one or more scientific publications. The responsible parties decide the publication forums. The first publication will include the results from the analysis of the primary outcome(s) that will include data from both Denmark and Sweden.

The investigators will inform the Sponsor in advance about any plans to publish or present data from the study. Any publications and presentations of the results (abstracts in journals or newspapers, oral presentations, etc.), either in whole or in part, by investigators or their representatives will require pre-submission review and approval by the Sponsor. The Sponsor has the right to comment on results and the interpretation thereof. Requests on changing the interpretation of the results or their presentation must be based on sound scientific reasons. The principal investigator is free not to take the comments of the Sponsor into account and, in the event of such a refusal, the Sponsor may only require that the presentation of the results be changed to delete confidential information. The Sponsor cannot unjustifiably delay the publication. In this particular study the commenting time for the Sponsor during the review rounds is agreed to be maximum of one month.

The principal investigator and the Sponsor are committed to ensuring that the authorship of all publications complies with the criteria defined by the International Committee of Medical Journal Editors (ICMJE). It is stated that each author should have participated sufficiently in the work to take public responsibility for the content. These conditions apply equally to external investigators and to the employees of the Sponsor.

12.3 Cost Results in Health-Economic Studies

After the completion of this non-interventional retrospective cohort study, results on treatment-associated costs will be used to assess the cost-effectiveness of rhFSH-alfa reference product, as summarized in Appendix 3 (Section 14.3.4).

13 References

1. Bassett R, Lispi M, Ceccarelli D, Grimaldi L, Mancinelli M, Martelli F, et al. Analytical identification of additional impurities in urinary-derived gonadotrophins. Reprod Biomed Online. 2009 Sep;19(3):300–13.

2. European Medicines Agency. GONAL-f [Internet]. 2018 [cited 2020 Jul 14]. Available from: https://www.ema.europa.eu/en/medicines/human/EPAR/gonal-f

3. Bassett RM, Driebergen R. Continued improvements in the quality and consistency of follitropin alfa, recombinant human FSH. Reprod Biomed Online. 2005 Feb;10(2):169–77.

4. Lunenfeld B, Bilger W, Longobardi S, Alam V, D'Hooghe T, Sunkara SK. The Development of Gonadotropins for Clinical Use in the Treatment of Infertility. Front Endocrinol. 2019;10:429.

5. Ulloa-Aguirre A, Timossi C, Méndez JP. Is there any physiological role for gonadotrophin oligosaccharide heterogeneity in humans? I. Gondatrophins are synthesized and released in multiple molecular forms. A matter of fact. Hum Reprod Oxf Engl. 2001 Apr;16(4):599–604.

6. Toegel S, Pabst M, Wu SQ, Grass J, Goldring MB, Chiari C, et al. Phenotype-related differential alpha-2,6- or alpha-2,3-sialylation of glycoprotein N-glycans in human chondrocytes. Osteoarthritis Cartilage. 2010 Feb;18(2):240–8.

7. Lunenfeld B. Gonadotropin stimulation: past, present and future. Reprod Med Biol. 2012 Jan;11(1):11–25.

8. Bosch E, Vidal C, Labarta E, Simon C, Remohi J, Pellicer A. Highly purified hMG versus recombinant FSH in ovarian hyperstimulation with GnRH antagonists--a randomized study. Hum Reprod Oxf Engl. 2008 Oct;23(10):2346–51.

9. Andersen AN, Devroey P, Arce J-C. Clinical outcome following stimulation with highly purified hMG or recombinant FSH in patients undergoing IVF: a randomized assessor-blind controlled trial. Hum Reprod Oxf Engl. 2006 Dec;21(12):3217–27.

10. Witz CA, Daftary GS, Doody KJ, Park JK, Seifu Y, Yankov VI, et al. Randomized, assessor-blinded trial comparing highly purified human menotropin and recombinant follicle-stimulating hormone in high responders undergoing intracytoplasmic sperm injection. Fertil Steril. 2020 Aug;114(2):321–30.

11. Strehler E, Abt M, El-Danasouri I, De Santo M, Sterzik K. Impact of recombinant follicle-stimulating hormone and human menopausal gonadotropins on in vitro fertilization outcome. Fertil Steril. 2001 Feb;75(2):332–6.

12. Bühler KF, Guedes S, Allignol A, Dâ□™Hooghe TM, Bilger W, Boutmy E, et al. Effectiveness of recombinant human follicle-stimulating hormone (r-hFSH) versus human menopausal Gonadotropin (u-hmg) in assisted reproductive technology (ART). A study based

on German Real-World data. In ASRM; 2019 [cited 2021 Mar 23]. Available from: https://asrm.confex.com/asrm/2019/meetingapp.cgi/Paper/3103

13. U.S. Food and Drug Administration (FDA). Biosimilar Development, Review, and Approval. 2019 Dec 3 [cited 2021 Mar 23]; Available from: https://www.fda.gov/drugs/biosimilars/biosimilar-development-review-and-approval

14. European Medicines Agency (EMA). Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues. :13.

15. ESHRE Reproductive Endocrinology Guideline Group. Ovarian Stimulation for IVF/ICSI. Guideline of the European Society of Human Reproduction and Embryology. 2019.

16. Zegers-Hochschild F, Adamson GD, Dyer S, Racowsky C, de Mouzon J, Sokol R, et al. The International Glossary on Infertility and Fertility Care, 2017. Fertil Steril. 2017;108(3):393–406.

17. Longobardi S, Mol B, Chua SJ, Wang R, Venetis C, Papsch R, et al. A systematic review and meta-analysis comparing Biosimilar versus Originator Recombinant Follitropin Alfa in women undergoing IVF/ICSI. In: ESHRE Annual Meeting Vienna 2019.

18. Mastrangeli R, Satwekar A, Cutillo F, Ciampolillo C, Palinsky W, Longobardi S. In-vivo biological activity and glycosylation analysis of a biosimilar recombinant human follicle-stimulating hormone product (Bemfola) compared with its reference medicinal product (GONAL-f). PloS One. 2017;12(9):e0184139.

19. Katz P, Showstack J, Smith JF, Nachtigall RD, Millstein SG, Wing H, et al. Costs of infertility treatment: Results from an 18-month prospective cohort study. Fertil Steril. 2011 Mar 1;95(3):915–21.

20. Smith ADAC, Tilling K, Nelson SM, Lawlor DA. Live-Birth Rate Associated With Repeat In Vitro Fertilization Treatment Cycles. JAMA. 2015 Dec 22;314(24):2654–62.

21. Chambers GM, Paul RC, Harris K, Fitzgerald O, Boothroyd CV, Rombauts L, et al. Assisted reproductive technology in Australia and New Zealand: cumulative live birth rates as measures of success. Med J Aust. 2017 Aug 7;207(3):114–8.

22. Stürmer T, Wyss R, Glynn RJ, Brookhart MA. Propensity scores for confounder adjustment when assessing the effects of medical interventions using nonexperimental study designs. J Intern Med. 2014 Jun;275(6):570–80.

23. Lee KH, Kim SH, Jee BC, Kim YJ, Suh CS, Kim KC, et al. Comparison of clinical characteristics between early and late patterns in hospitalized patients with ovarian hyperstimulation syndrome. Fertil Steril. 2010 May 1;93(7):2274–80.

24. Retsinformation. Vejledning om sundhedspersoners og vævscentres virksomhed og forpligtelser i forbindelse med assisteret reproduktion [Internet]. 2015 [cited 2020 Oct 22].

Available from: https://www.retsinformation.dk/eli/retsinfo/2015/9351#ide681ffa8-817c-4e1e-96ff-5ce8c2fa0a2a

25. Retsinformation. Vejledning om autoriserede sundhedspersoners og vævscentres virksomhed og forpligtelser i forbindelse med kunstig befrugtning [Internet]. [cited 2021 Apr 15]. Available from: https://www.retsinformation.dk/eli/retsinfo/2012/9628

26. Retsinformation. Vejledning om kunstig befrugtning og anden reproduktionsfremmende behandling Til landets læger [Internet]. [cited 2021 Apr 15]. Available from: https://www.retsinformation.dk/eli/retsinfo/1997/15057

27. Christensen MW, Kesmodel US, Christensen K, Kirkegaard K, Ingerslev HJ. Early ovarian ageing: is a low number of oocytes harvested in young women associated with an earlier and increased risk of age-related diseases? Hum Reprod Oxf Engl. 2020 Oct 1;35(10):2375–90.

28. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. Scand J Public Health. 2011 Jul;39(7 Suppl):30–3.

29. Morten Schmidt, Sigrun Alba Johannesdottir Schmidt, Jakob Lynge Sandegaard, Vera Ehrenstein, Lars Pedersen, and Henrik Toft Sørensen. The Danish National Patient Registry: a review of content, data quality, and research potential [Internet]. [cited 2020 Oct 27]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4655913/

30. Bliddal M, Broe A, Pottegård A, Olsen J, Langhoff-Roos J. The Danish Medical Birth Register. Eur J Epidemiol. 2018 Jan;33(1):27–36.

31. Kildemoes HW, Sørensen HT, Hallas J. The Danish National Prescription Registry. Scand J Public Health. 2011 Jul;39(7 Suppl):38–41.

32. Pottegård A, Schmidt SAJ, Wallach-Kildemoes H, Sørensen HT, Hallas J, Schmidt M. Data Resource Profile: The Danish National Prescription Registry. Int J Epidemiol. 2017 Jun 1;46(3):798–798f.

33. Nationellt kvalitetsregister för assisterad befruktning (Q-IVF). Fertilitetsbehandlingar i Sverige. Årsrapport 2020.

34. Nationellt kvalitetsregister för assisterad befruktning (Q-IVF). Q-IVF — Behandlingsresultat t.o.m. 2013 [Internet]. [cited 2021 Mar 26]. Available from: https://www.medscinet.com/qivf/behandlingsresultat-tom-2013.aspx

35. Nationellt kvalitetsregister för assisterad befruktning (Q-IVF). Q-IVF — Årsrapporter [Internet]. [cited 2021 Mar 26]. Available from: https://www.medscinet.com/qivf/arsrapporter.aspx

36. Dansk Fertilitetsselskab. DATAARK PR ÅR [Internet]. [cited 2021 Mar 26]. Available from: https://fertilitetsselskab.dk/dataark-pr-aar/

37. Unpublished report [Feasibility Study for RWE Study in the Nordics: GONAL-f vs. biosimilars and hMG. 2020].

38. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med. 2015 Dec 10;34(28):3661–79.

39. Malizia BA, Hacker MR, Penzias AS. Cumulative live-birth rates after in vitro fertilization. N Engl J Med. 2009 Jan 15;360(3):236–43.

40. Malchau SS, Henningsen AA, Forman J, Loft A, Nyboe Andersen A, Pinborg A. Cumulative live birth rate prognosis based on the number of aspirated oocytes in previous ART cycles. Hum Reprod. 2019 Jan 1;34(1):171–80.

41. Magnusson Å, Källen K, Thurin-Kjellberg A, Bergh C. The number of oocytes retrieved during IVF: a balance between efficacy and safety. Hum Reprod Oxf Engl. 2018 Jan 1;33(1):58–64.

42. European Medicines Agency (EMA). The ENcePP Code Of Conduct For Scientific Independence And Transparency In The Conduct Of Pharmacoepidemiological And Pharmacovigilance Studies.

43. Public Policy Committee, International Society of Pharmacoepidemiology. Guidelines for good pharmacoepidemiology practice (GPP). Pharmacoepidemiol Drug Saf. 2016 Jan;25(1):2–10.

44. Medstat. Sundhedsdatastyrelsen - Statistikker [Internet]. [cited 2021 Mar 23]. Available from: https://medstat.dk/en

45. Langhoff- Roos J, Krebs L, Klungsøyr K, Bjarnadottir RI, Källén K, Tapper A-M, et al. The Nordic medical birth registers – a potential goldmine for clinical research. Acta Obstet Gynecol Scand. 2014;93(2):132–7.

46. Socialstyrelsen. Bortfall och kvalitet om medicinska födelseregistret [Internet]. Socialstyrelsen. [cited 2021 Mar 23]. Available from: https://www.socialstyrelsen.se/statistik-och-data/register/alla-register/medicinska-fodelseregistret/bortfall-och-kvalitet/

47. Nationellt kvalitetsregister för assisterad befruktning (Q-IVF). Fertilitetsbehandlingar i Sverige. Årsrapport 2019.

48. Marklund A, Lundberg FE, Eloranta S, Hedayati E, Pettersson K, Rodriguez-Wallberg KA. Reproductive Outcomes After Breast Cancer in Women With vs Without Fertility Preservation. JAMA Oncol. 2021 Jan 1;7(1):86–91.

49. Rodriguez-Wallberg KA, Lundberg FE, Ekberg S, Johansson ALV, Ludvigsson JF, Almqvist C, et al. Mortality from infancy to adolescence in singleton children conceived from assisted reproductive techniques versus naturally conceived singletons in Sweden. Fertil Steril. 2020 Mar;113(3):524–32.

50. Johansen G, Dahm-Kähler P, Staf C, Flöter Rådestad A, Rodriguez-Wallberg KA. Reproductive and obstetrical outcomes with the overall survival of fertile-age women treated with fertility-sparing surgery for borderline ovarian tumors in Sweden: a prospective nationwide population-based study. Fertil Steril. 2021 Jan;115(1):157–63.

51. Cesta CE, Viktorin A, Olsson H, Johansson V, Sjölander A, Bergh C, et al. Depression, anxiety, and antidepressant treatment in women: association with in vitro fertilization outcome. Fertil Steril. 2016 Jun;105(6):1594-1602.e3.

52. Andersen AN, Westergaard HB, Olsen J. The Danish in vitro fertilisation (IVF) register. Dan Med Bull. 1999 Sep;46(4):357–60.

53. Kristensen J, Langhoff-Roos J, Skovgaard LT, Kristensen FB. Validation of the danish birth registration. J Clin Epidemiol. 1996 Aug 1;49(8):893–7.

54. Larsen MD, Jensen DM, Fedder J, Jølving LR, Nørgård BM. Live-born children after assisted reproduction in women with type 1 diabetes and type 2 diabetes: a nationwide cohort study. Diabetologia. 2020 Sep 1;63(9):1736–44.

55. Blenstrup LT, Knudsen LB. Danish registers on aspects of reproduction. Scand J Public Health. 2011 Jul;39(7 Suppl):79–82.

56. National Board of Health, Denmark. IVF treatments in Denmark 1998–2005 [in Danish] Nye tal fra Sundhedsstyrelsen 2007;14. Copenhagen.

57. Hansen AT, Kesmodel US, Juul S, Hvas AM. Increased venous thrombosis incidence in pregnancies after in vitro fertilization. Hum Reprod Oxf Engl. 2014 Mar;29(3):611–7.

58. Hargreave M, Jensen A, Hansen MK, Dehlendorff C, Winther JF, Schmiegelow K, et al. Association Between Fertility Treatment and Cancer Risk in Children. JAMA. 2019 Dec 10;322(22):2203–10.

59. Socialstyrelsen. Statistik om aborter [Internet]. Socialstyrelsen. [cited 2021 Mar 22]. Available from: https://www.socialstyrelsen.se/statistik-och-data/statistik/statistikamnen/aborter/

60. Brodin T, Hadziosmanovic N, Berglund L, Olovsson M, Holte J. Antimüllerian hormone levels are strongly associated with live-birth rates after assisted reproduction. J Clin Endocrinol Metab. 2013 Mar;98(3):1107–14.

61. Holte J, Brodin T, Berglund L, Hadziosmanovic N, Olovsson M, Bergh T. Antral follicle counts are strongly associated with live-birth rates after assisted reproduction, with superior treatment outcome in women with polycystic ovaries. Fertil Steril. 2011 Sep;96(3):594–9.

62. European Medicines Agency, Heads of Medicines Agencies. Guideline on good pharmacovigilance practices (GVP)-Module VI–Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2) [Internet]. 2017 [cited 2020 Jun 16]. Available from: https://www.ema.europa.eu/en/documents/regulatory-

procedural-guideline/guideline-good-pharmacovigilance-practices-gvp-module-vi-collection-management-submission-reports_en.pdf

14 Appendices

14.1 List of Stand-Alone Documents

None.

14.2 ENCePP Checklist for Study Protocols

Doc.Ref. EMA/540136/2009

European Network of Centres for Pharmacoepidemiology and Pharmacovigilance

ENCePP Checklist for Study Protocols (Revision 4)

Adopted by the ENCePP Steering Group on 15/10/2018

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) welcomes innovative designs and new methods of research. This Checklist has been developed by ENCePP to stimulate consideration of important principles when designing and writing a pharmacoepidemiological or pharmacovigilance study protocol. The Checklist is intended to promote the quality of such studies, not their uniformity. The user is also referred to the ENCePP Guide on Methodological Standards in Pharmacoepidemiology, which reviews and gives direct electronic access to guidance for research in pharmacoepidemiology and pharmacovigilance.

For each question of the Checklist, the investigator should indicate whether or not it has been addressed in the study protocol. If the answer is "Yes", the section number of the protocol where this issue has been discussed should be specified. It is possible that some questions do not apply to a particular study (for example, in the case of an innovative study design). In this case, the answer 'N/A' (Not Applicable) can be checked and the "Comments" field included for each section should be used to explain why. The "Comments" field can also be used to elaborate on a "No" answer.

This Checklist should be included as an Annex by marketing authorisation holders when submitting the protocol of a non-interventional post-authorisation safety study (PASS) to a regulatory authority (see the Guidance on the format and content of the protocol of non-interventional post-authorisation safety studies). The Checklist is a supporting document and does not replace the format of the protocol for PASS presented in the Guidance and Module VIII of the Good pharmacovigilance practices (GVP).

Study title: A non-interventional register-based comparative effectiveness study of rhFSH-alfa reference product vs. highly purified human menopausal gonadotropin or rhFSH-alfa biosimilar products for ovarian stimulation in in vitro fertilization or intracytoplasmic sperm injection treatment in Denmark and Sweden – The Nordic Follitropin Alfa Comparative Effectiveness (NORD-FACE) Study

EU PAS Register® number:
Study reference number (if applicable):

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				6
1.1.1 Start of data collection[1]	☒	☐	☐	
1.1.2 End of data collection[2]	☒	☐	☐	
1.1.3 Progress report(s)	☐	☐	☒	
1.1.4 Interim report(s)	☒	☐	☐	

[1] Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.
[2] Date from which the analytical dataset is completely available.

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	
1.1.6 Final report of study results.	☒	☐	☐	

Comments:

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1 Does the formulation of the research question and objectives clearly explain:	☐	☐	☐	7, 8
2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	
2.1.2 The objective(s) of the study?	☒	☐	☐	
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	
2.1.4 Which hypothesis(-es) is (are) to be tested?	☒	☐	☐	
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

Section 3: Study design	**Yes**	**No**	**N/A**	**Section Number**
3.1 Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	☒	☐	☐	9.1.1
3.2 Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	9.1.1
3.3 Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	9.7.3
3.4 Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☒	☐	☐	9.7.3
3.5 Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☒	☐	☐	11

Comments:

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.1 Is the source population described?	☒	☐	☐	9.2.1
4.2 Is the planned study population defined in terms of:				

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.2.1 Study time period	☒	☐	☐	9.1.1, 9.2.3
4.2.2 Age and sex	☒	☐	☐	9.2.1
4.2.3 Country of origin	☒	☐	☐	9.1.1, 9.2.1
4.2.4 Disease/indication	☒	☐	☐	9.1.1, 9.2.1
4.2.5 Duration of follow-up	☒	☐	☐	9.1.1, 9.2.3
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	9.2.1

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.1 Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	9.2.2, 9.3.1
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☒	☐	☐	9.2.2, 9.3.1, 9.7.3.9, 9.9.2
5.3 Is exposure categorised according to time windows?	☒	☐	☐	
5.4 Is intensity of exposure addressed? (e.g. dose, duration)	☐	☒	☐	
5.5 Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☒	☐	
5.6 Is (are) (an) appropriate comparator(s) identified?	☒	☐	☐	

Comments:

Section 6: Outcome definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
6.1 Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☒	☐	☐	9.1.2
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	9.1.2, 9.7.3, 14.3.1.1

Section 6: Outcome definition and measurement	Yes	No	N/A	Section Number
6.3 Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☒	☐	☐	9.9.2
6.4 Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☐	☐	☒	

Comments:

Section 7: Bias	Yes	No	N/A	Section Number
7.1 Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☒	☐	☐	9.7.3.1
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☒	☐	☐	9.9.1
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	9.9.2

Comments:

Section 8: Effect measure modification	Yes	No	N/A	Section Number
8.1 Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☒	☐	☐	9.7.3.6, 9.7.3.9

Comments:

Section 9: Data sources	Yes	No	N/A	Section Number
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	9.3.1
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	9.1.2, 14.3.1.1
9.1.3 Covariates and other characteristics?	☒	☐	☐	9.3.2, 14.3.1.3
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	9.3.1, 9.4

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	9.1.2, 9.4
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	9.3.2, 9.4
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	9.3.1, 14.3.1.3
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	9.1.2, 14.3.1.1
9.3.3 Covariates and other characteristics?	☒	☐	☐	9.3.2, 14.3.1.3
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	9.1.1, 9.4

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	9.7
10.2 Is study size and/or statistical precision estimated?	☒	☐	☐	9.5
10.3 Are descriptive analyses included?	☒	☐	☐	9.7.3.2
10.4 Are stratified analyses included?	☒	☐	☐	9.7.3.6
10.5 Does the plan describe methods for analytic control of confounding?	☒	☐	☐	9.7.3
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☒	☐	☐	9.7.3.3, 9.7.3.9
10.7 Does the plan describe methods for handling missing data?	☒	☐	☐	9.7.2
10.8 Are relevant sensitivity analyses described?	☒	☐	☐	9.7.3.9

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	9.6
11.2 Are methods of quality assurance described?	☒	☐	☐	9.10.2
11.3 Is there a system in place for independent review of study results?	☒	☐	☐	9.8, 12

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☒	☐	☐	9.9.1
12.1.2 Information bias?	☒	☐	☐	9.9.2
12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☒	☐	☐	9.9.4
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☒	☐	☐	9.5, 9.9.2

Comments:

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	9.10.1
13.2 Has any outcome of an ethical review procedure been addressed?	☐	☐	☒	
13.3 Have data protection requirements been described?	☒	☐	☐	9.6, 10.1

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	5

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	12.1
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	12.2

Comments:

Name of the main author of the protocol: PPD

Date: 29/April/2021

Signature: PPD

14.3 Additional Information

All code lists included below may be subject to amendment or revision during study conduct.

14.3.1 Outcome variables (Section 9.1.2): Diagnosis and Procedure Codes for Definition and Databases for Identification

Outcome	Code for identification[+]	Comment	Source for identification
Primary outcome			
Live birth	N/A		DMBR, SMBR, Q-IVF
Secondary outcomes			
Clinical pregnancy	ICD-10: O00 NOMESCO procedure codes: MCB 00, LAA 20, LBC*	Intrauterine pregnancy ascertained in the DIVF and Q-IVF. ICD-10: Ectopic pregnancy NOMESCO procedures: laparotomy and delivery of extrauterine fetus, percutaneous or transvaginal injection into ovarian pregnancy, tube conserving operations for tubal pregnancy	DIVF, Q-IVF, DNPR, SNPR
Ongoing pregnancy	ICD-10: O03, O04, O05, O06[#] NOMESCO procedure codes: LCH*	Delivery and gestational age at delivery ascertained in the DMBR and SMBR. For Sweden, miscarriage and induced abortion† will be ascertained in the SNPR and Q-IVF. ICD-10: spontaneous abortion, medical abortion, other abortion, unspecified abortion NOMESCO procedures: procedures related to termination of pregnancy	DMBR, SMBR, DNPR, SNPR, Q-IVF
Oocytes retrieved	N/A	Ascertained using information recorded in the DIVF and Q-IVF.	DIVF, QIVF
Embryos transferred			
Embryos cryopreserved			
Implantation			
Pregnancy loss	ICD-10: O03, O04, O05, O06[#] NOMESCO procedure codes: LCH*	ICD-10: spontaneous abortion, medical abortion, other abortion, unspecified abortion NOMESCO procedures: procedures related to termination of pregnancy	DNPR, SNPR, Q-IVF

Outcome	Code for identification[+]	Comment	Source for identification
		For Sweden, miscarriage and induced abortion† will be ascertained in the SNPR and Q-IVF.	
Multiple pregnancy		Ascertained in the DIVF and Q-IVF based on ultrasonographic visualization of two or more gestational sacs. To capture multiple pregnancy with monozygotic twins, which share the same gestational sac, an algorithm will be defined. The algorithm will consider intrauterine pregnancy loss (miscarriage or induced abortion) before 22 completed weeks of gestational age and live or stillbirth after 22 completed weeks of gestational age. Thus, the algorithm accounts for reduction of multiple pregnancies. This will be particularly important for Denmark, as the DIVF does not record pregnancy outcome. For Sweden, however, the Q-IVF captures pregnancy outcome for each fetus, in case of multiple pregnancy. The algorithm will be detailed in the SAP.	DIVF, Q-IVF, DNPR, SNPR, DMBR, SMBR
	ICD-10: O03, O04, O05, O06[#]	ICD-10: spontaneous abortion, medical abortion, other abortion, unspecified abortion	
	NOMESCO procedure codes: LCH*	NOMESCO procedures: procedures related to termination of pregnancy	
Cycle cancellation	N/A	Ascertained using information recorded in the DIVF and Q-IVF.	DIVF, QIVF
OHSS	ICD-10: N98.1	ICD-10: Hyperstimulation of ovaries For Sweden, OHSS is recorded in the Q-IVF as a) requiring inpatient hospitalization with/without ascites drainage, or b) requiring ascites drainage in outpatient specialized care. In both countries, the national patient registers and the IVF registers will be used to ascertain information on OHSS.	DNPR, SNPR, Q-IVF
Treatment-associated costs	To be defined	Relevant ATC, procedure, ICD-10, and DRG codes (including associated costs/weights) and cost listings will be used to ascertain costs. These codes will be further detailed in the SAP.	RMPS, SPDR, DIVF, Q-IVF, DMBR, SMBR, DNPR, SNPR, cost listings‡

ATC: Anatomical Therapeutic Chemical; DIVF: Danish In-vitro Fertilization Register; DMBR: Danish Medical Birth Register; DNPR: Danish National Patient Register; DRG: diagnosis-related group; ICD-10: International Classification of Diseases, 10th revision; IVF: in vitro fertilization; N/A: not applicable; NOMESCO: Nordic Medico-Statistical Committee; Q-IVF: Swedish National Quality Registry of Assisted Reproduction; RMPS: Danish Register of Medicinal Product Statistics; SAP: statistical analysis plan; SMBR: Swedish Medical Birth Register; SNPR: Swedish National Patient Register; SPDR: Swedish Prescribed Drug Register.

Notes:

[+]Country-specific adaptations may be made at the data permit stage. Conditions defined by ICD-10 codes are considered present if there was a record of a primary or secondary diagnosis in the national (patient) registers (outpatient or inpatient care). Similarly, procedures defined by NOMESCO codes are considered to have been performed if there was a record in the national patient registers (outpatient or inpatient care). Conditions ascertained in the IVF registers (not using ICD-10 codes) are considered present if there was a relevant record in those registers.

*All ICD-10 or procedure subcodes are included.

[#]The DNPR additionally records the gestational age at time of abortion.

†Prior to October 2016, induced abortions were not recorded in the SNPR.

‡If costs of healthcare resource utilization are unavailable at the individual-level in the register data, available national and regional cost listings will be used for unit costs for healthcare resource utilization.

14.3.2 Censoring Criteria (Section 9.2.3): Diagnosis and Procedure Codes for Definition

Criterion	Diagnosis/procedure code[+]	Comment
Death (any cause)	N/A	
Hysterectomy	ICD-10: N99.3, O82.2, Z90.7 NOMESCO procedure codes: LCC*, LCD*, LCH 20	ICD-10: Prolapse of vaginal vault after hysterectomy, delivery by caesarean hysterectomy, acquired absence of genital organ (hysterectomy) NOMESCO procedures: procedures related to partial excision of uterus or total excision of uterus
Bilateral oophorectomy	ICD-10: Z90.7 NOMESCO procedure codes: LAE 20, LAE 21, LAF 10, LAF 11, LAF 30	ICD-10: Acquired absence of genital organ(s) NOMESCO procedures: bilateral oophorectomy, bilateral laparoscopic oophorectomy, bilateral salpingo-oophorectomy, laparoscopic bilateral salpingo-oophorectomy, bilateral transvaginal salpingo-oophorectomy
Breast cancer[#]	ICD-10: C50*	Malignant neoplasm of breast
Ovarian cancer[#]	ICD-10: C56*	Malignant neoplasm of ovary
Uterine cancer[#]	ICD-10: C53*, C54*, C55*	Malignant neoplasm of cervix uteri, malignant neoplasm of corpus uteri, malignant neoplasm of uterus, part unspecified

ATC: Anatomical Therapeutic Chemical; ICD-10: International Classification of Diseases, 10th revision; NOMESCO: Nordic Medico-Statistical Committee.

Notes:

[+]Country-specific adaptations may be made at the data permit stage. Conditions defined by ICD-10 codes are considered present if there was a record of a primary or secondary diagnosis in the national (patient) registers (outpatient or inpatient care)]. Similarly, procedures defined by NOMESCO codes are considered to have been performed if there was a record in the national patient registers (outpatient or inpatient care).

*All ICD-10 or procedure subcodes are included.

[#] Identification in the national patient registers, as opposed to the cancer registers, is considered sufficient for the intended use in this study.

14.3.3 Exposure (Section 9.3.1) and Other Variables (Section 9.3.2): Codes for Definition

Variable	Code[+]	Comment
Type of gonadotropin used for COS, including study and non-study drugs (ATC codes)		
Follitropin alfa	G03GA05	Distinction between rhFSH-alfa reference and biosimilar products possible using brand name
Human menopausal gonadotropin	G03GA02	
Urofollitropin	G03GA04	
Follitropin beta	G03GA06	
Corifollitropin alfa	G03GA09	
Combinations (of recombinant hormones)	G03GA30	
GnRH drugs (ATC codes)		
Nafarelin	H01CA02	
Ganirelix	H01CC01	
Cetrorelix	H01CC02	
Buserelin	L02AE01	
Drugs used for ovulation triggering (ATC codes)		
Chorionic gonadotropin (human urinary)	G03GA01	
Choriogonadotropin alfa (human recombinant)	G03GA08	
Drugs used for luteal phase support (ATC codes)		
Progesterone	G03DA04	Dosage form determined by other information recorded with the dispensed prescription
Fertility-related medical history and other clinical characteristics (ATC, ICD-10 and procedure codes)		
Obesity	ICD-10: E66*	
Bariatric surgery	NOMESCO procedure code: JDF*	Bariatric operations on stomach: Operations for morbid obesity and intestinal bypass operations
Type 1 diabetes mellitus	ICD-10: E10*	
Type 2 diabetes mellitus	ICD-10: E11*	
Thromboembolic events	ICD-10: I21*, I26*, I63*, I65*, I66*, I67.6, I74*, I80.1, I80.2, I80.3, I80.8, I80.9, I81*, I82*, H34*, K55.0, O22.3, O22.5, O22.8, O22.9, N28.0	Pulmonary embolism, cerebral infarction, thromboembolic disease unspecified, thromboflebitis (excluding superficial), deep thrombosis in pregnancy, venous complication in pregnancy, other
Thrombosis prophylaxis	ATC: B01A*	Antithrombotic agents, including vitamin K antagonists, heparin, platelet aggregation inhibitors, enzymes, direct thrombin inhibitors, direct factor Xa inhibitors, other
Antidiabetic drugs	ATC: A10*	Insulins and analogues, blood glucose lowering drugs, other drugs used in diabetes
Any cancer	ICD-10: C00-C97*	Malignant neoplasms
Porphyria	ICD-10: E80*	Disorders of porphyrin and bilirubin metabolism

Variable	Code+	Comment
Amenorrhea or oligomenorrhea	ICD-10: N91*	Absent, scanty and rare menstruation, excluding ovarian dysfunction
PCOS	ICD-10: E28.2	
Endometriosis	ICD-10: N80*	
Pituitary insufficiency	ICD-10: E23*	
Female infertility	ICD-10: E28.3, E28.8, E28.9, N97*	Primary ovarian failure, other specified ovarian dysfunction, ovarian dysfunction unspecified, female infertility (excluding female infertility associated with male factors)
Thyroid disease	ICD-10: E00*-E07*	Disorders of thyroid gland
Parathyroid diseases	ICD-10: E21*, E20*	Hypoparathyroidism, hyperparathyroidism
Disorders of adrenal gland	ICD-10: E24.0, E26*, E27* (excluding E27.2, E27.3, E27.5)	Pituitary-dependent Cushing disease, hyperaldosteronism and other disorders of adrenal gland, including Addison disease
Polyglandular dysfunction	ICD-10: E31*	
Polyarthritis nodosa	ICD-10: M30.0-M30.2	
Rheumatoid arthritis	ICD-10: M05*-M14*	
SLE and other systemic connectivity tissue disorders	ICD-10: M32.1, M32.8, M32.9, M33*, M34* (excluding M34.2), M35.0	Includes dermatopolymyositis, systemic sclerosis and Sjögren's syndrome
Inflammatory bowel disease	ICD-10: K50*, K51*	Crohn disease, ulcerative colitis
Epilepsy	ICD-10: G40*, G41*	Epilepsy, status epilepticus
Multiple sclerosis	ICD-10: G35*	
Hypertension	ICD-10: I10*-I15*	Hypertensive diseases
Ischemic heart diseases	ICD-10: I20*-I25*	
Chronic heart diseases	ICD-10: I27*, I31*, I34*-I37*, I39*, I42*, I52*	Other pulmonary heart diseases, other diseases of pericardium, heart valve disorders, cardiomyopathy, other heart disorders in diseases classified elsewhere
Cerebrovascular diseases	ICD-10: I60*-I69*	
Atherosclerosis	ICD-10: I70*	
Coagulation disorders	ICD-10: D66*-D68*	Hereditary factor VIII deficiency, hereditary factor IX deficiency, other coagulation defects
Chronic lung disease	ICD-10: J40*-J47*, J60*-J70*, J84*	Chronic lower respiratory diseases, lung diseases due to external agents, other interstitial pulmonary diseases
HIV	ICD-10: B20*-B24*	
Mood disorders	ICD-10: F30*-F39*	
Schizophrenia, schizotypal and delusional disorders	ICD-10: F20*-F29*	
Neurotic, stress-related and somatoform disorders, specific personality disorders	ICD-10: F40*-F48*, F60*	
Pervasive developmental disorders	F84* (excluding F84.2, F84.3, F84.4)	
Antidepressants	ATC: N06A	

Variable	Code[+]	Comment
Antipsychotics	ATC: N05A	
Anxiolytics	ATC: N05B	
Hypnotics and sedatives	ATC: N05C	
Medical history, laboratory test results[#]		
Anti-Müllerian hormone	NPU27675 NPU27385	
FSH	NPU04014 NPU02072 NPU02073	
Thyrotropin	DNK05280 NPU04026 NPU04199 NPU10347 NPU10374 NPU04200 NPU08717 NPU04202 NPU19580 NPU29633 NPU03577 NPU27547 NPU28066 NPU28065 NPU03751	
Prolactin	NPU21694 NPU04022 NPU04021 NPU18247 NPU03252 NPU26171 NPU26172 NPU26173 NPU26174 NPU22291 NPU22292 NPU26175 NPU22293 NPU26176 NPU26177 NPU26178 NPU19897 NPU10458 NPU10460 NPU10682	

ATC: Anatomical Therapeutic Chemical; COS: controlled ovarian stimulation; FSH: follicle-stimulating hormone; ICD-10: International Classification of Diseases, 10th revision; HIV: human immunodeficiency virus; IVF: in vitro fertilization; NOMESCO: Nordic Medico-Statistical Committee; NPU: Nomenclature for Property and Unit; PCOS: polycystic ovarian syndrome; rhFSH: recombinant human follicle-stimulating hormone; SAP: statistical analysis plan; SLE: systemic lupus erythematosus.

Notes:

[+]Country-specific adaptations may be made at the data permit stage. Conditions defined by ICD-10 codes are considered present if there was a record of a primary or secondary diagnosis in the national (patient) registers (outpatient or inpatient care). Similarly, procedures defined by NOMESCO codes are considered to have been

performed if there was a record in the national patient registers (outpatient or inpatient care). Based on dispensing records in the prescribed drug registers for drugs defined by the listed ATC codes, it is assumed a given drug was used.

*All ATC, ICD-10, or procedure subcodes are included.

#Available for Denmark only.

14.3.4 Summary of the Cost-effectiveness Model

Cost data collected in this study will be fed into a cost-effectiveness model developed by PPD for the Sponsor, to assess the cost effectiveness of GONAL-f®. An existing cost-effectiveness model using in other countries will be updated and adapted to reflect the cost perspectives in the Danish and Swedish settings. The cost-effectiveness analysis aims to support the value story and benefits of rhFSH-alfa reference product in comparison with HP-hMG, rhFSH-alfa biosimilar products and other rhFSH products in terms of improved live birth rates and cost-per-outcome ratio (cost per live birth), as well as potential savings resulting from less occurrence of adverse events.

The model uses the decision-tree framework designed specifically for IVF/ICSI treatments. Thus, it models the cost per live birth associated with rhFSH-alfa reference product, GONAL-f®, against each comparator drug (HP-hMG and rhFSH-alfa biosimilar products in this study). In addition, the model estimates the incremental cost-effectiveness ratio (ICER) for rhFSH-alfa reference product against the comparator drugs of this study. The decision-tree based model uses estimated probabilities of treatment procedures and outcomes based on observed clinical data, costs of treatments, pricing of the study drugs as well as adverse event (i.e., OHSS) costs as input parameters to obtain the average cost associated with achieved live birth and the ICER. In addition, the model contains a possibility to add additional relevant costs.

As part of the planned cost-effectiveness study, the results of this non-interventional study will be complemented with clinical expert opinion, used to ensure that cost inputs are representative of the actual costs of treatment procedures and pricing of the study drugs within each study country.

The model will be provided in Microsoft Excel 365 and delivered as a Microsoft Excel Binary Worksheet (.xlsb) file, which is macro-enabled in order to run properly.

14.4 Signature Pages and Responsible Persons for the Study

Signature Page – Protocol Lead

Study Title:	A non-interventional register-based comparative effectiveness study of rhFSH-alfa reference product vs. highly purified human menopausal gonadotropin or rhFSH-alfa biosimilar products for ovarian stimulation in *in vitro* fertilization or intracytoplasmic sperm injection treatment in Denmark and Sweden, 2010-2020 – The Nordic Follitropin Alfa Comparative Effectiveness (NORD-FACE) Study
Study Protocol Date / Version:	27 May 2021 / 1.0

Protocol Lead responsible for designing the non-interventional study:

I approve the design of the non-interventional study:

PPD _____ May 27, 2021 _____

Signature Date of Signature

Name, academic degree:	PPD
Function / Title:	PPD
Institution:	EMD Serono Research & Development Institute, Inc. an affiliate of Merck KGaA, Darmstadt, Germany
Address:	45A Middlesex Turnpike Billerica, MA, 01821, USA
Telephone number:	PPD
Fax number:	Not applicable
E-mail address:	PPD

Signature Page – Principal-in-Charge, Coordinating Center

Study Title	A non-interventional register-based comparative effectiveness study of rhFSH-alfa reference product vs. highly purified human menopausal gonadotropin or rhFSH-alfa biosimilar products for ovarian stimulation in *in vitro* fertilization or intracytoplasmic sperm injection treatment in Denmark and Sweden, 2010-2020 – The Nordic Follitropin Alfa Comparative Effectiveness (NORD-FACE) Study
Study Protocol Date / Version	27 May 2021 / 1.0
Coordinating Center	PPD
Principal-In-Charge	PPD

I approve the design of the non-interventional study and I understand and will conduct the study according to the study protocol, any approved protocol amendments, Good Pharmacoepidemiology Practices (GPP) and all applicable Health Authority requirements and national laws.

PPD _____ May 27, 2021

Date of Signature

Name, academic degree:	PPD
Function / Title:	PPD
Institution:	PPD
Address:	PPD
Telephone number:	PPD
Fax number:	Not applicable
E-mail address:	PPD

Signature Page – Principal Investigator

Study Title	A non-interventional register-based comparative effectiveness study of rhFSH-alfa reference product vs. highly purified human menopausal gonadotropin or rhFSH-alfa biosimilar products for ovarian stimulation in *in vitro* fertilization or intracytoplasmic sperm injection treatment in Denmark and Sweden, 2010-2020 – The Nordic Follitropin Alfa Comparative Effectiveness (NORD-FACE) Study
Study Protocol Date / Version	27 May 2021 / 1.0
Coordinating Center	PPD
Principal Investigator	PPD

I approve the design of the non-interventional study and I understand and will conduct the study according to the study protocol, any approved protocol amendments, Good Pharmacoepidemiology Practices (GPP) and all applicable Health Authority requirements and national laws.

PPD

May 27, 2021

Date of Signature

Name, academic degree:	PPD
Function / Title:	PPD
Institution:	PPD
Address:	PPD
Telephone number:	PPD
Fax number:	Not applicable
E-mail address:	PPD PPD

Signature Page – Lead Epidemiologist and Protocol Lead, Coordinating Center

Study Title	A non-interventional register-based comparative effectiveness study of rhFSH-alfa reference product vs. highly purified human menopausal gonadotropin or rhFSH-alfa biosimilar products for ovarian stimulation in *in vitro* fertilization or intracytoplasmic sperm injection treatment in Denmark and Sweden, 2010-2020 – The Nordic Follitropin Alfa Comparative Effectiveness (NORD-FACE) Study
Study Protocol Date / Version	27 May 2021 / 1.0
Coordinating Center	PPD
Lead Epidemiologist, Protocol Lead	PPD

I approve the design of the non-interventional study:

PPD _____ May 28, 2021 _____

Signature Date of Signature

Name, academic degree:	PPD
Function / Title:	PPD
Institution:	PPD
Address:	PPD
Telephone number:	PPD
Fax number:	Not applicable
E-mail address:	PPD

Sponsor Responsible Persons not Named on the Cover Page

Name, academic degree:	PPD
Function / Title:	PPD
Institution:	Merck Healthcare KGaA
Address:	Frankfurter Str. 250 64293 Darmstadt, Germany
Telephone number:	PPD
E-mail address:	PPD

Name, academic degree:	PPD
Function / Title:	PPD
Institution:	EMD Serono Research & Development Institute, Inc. an affiliate of Merck KGaA, Darmstadt, Germany
Address:	45A Middlesex Turnpike Billerica, MA, 01821, USA
Telephone number:	PPD
E-mail address:	PPD