

Cost-effectiveness of ovarian stimulation agents for IUI in couples with unexplained subfertility

R. van Eekelen^{1,*}, R. Wang², N.A. Danhof¹, F. Mol¹, M. Mochtar¹, B.W. Mol², and M. van Wely¹

¹Centre for Reproductive Medicine, Amsterdam UMC, Location Academic Medical Centre, 1105 AZ Amsterdam, the Netherlands

²Department of Obstetrics and Gynaecology, Monash University, Clayton, VIC 3168, Australia

*Correspondence address. Centre for Reproductive Medicine, Amsterdam UMC, location Academic Medical Centre, Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands. E-mail: r.vaneekelen@amsterdamumc.nl

Submitted on July 31, 2020; resubmitted on December 21, 2020; editorial decision on January 11, 2021

STUDY QUESTION: Which agent for ovarian stimulation (OS) is the most cost-effective option in terms of net benefit for couples with unexplained subfertility undergoing IUI?

SUMMARY ANSWER: In settings where a live birth is valued at €3000 or less, between €3000 and €55 000 and above €55 000, clomiphene citrate (CC), Letrozole and gonadotrophins were the most cost-effective option in terms of net benefit, respectively.

WHAT IS KNOWN ALREADY: IUI-OS is a common first-line treatment for couples with unexplained subfertility and its increased uptake over the past decades and related personal or reimbursed costs are pressing concerns to patients and health service providers. However, there is no consensus on a protocol for conducting IUI-OS, with differences between countries, clinics and settings in the number of cycles, success rates, the agent for OS and the maximum number of dominant follicles in order to minimise the risk of a multiple pregnancy. In view of this uncertainty and the association with costs, guidance is needed on the cost-effectiveness of OS agents for IUI-OS.

STUDY DESIGN, SIZE, DURATION: We developed a decision-analytic model based on a decision tree that follows couples with unexplained subfertility from the start of IUI-OS to a protocolled maximum of six cycles, assuming couples receive four cycles on average within one year. We chose the societal perspective, which coincides with other perspectives such as that from health care providers, as the treatments are identical except for the stimulation agent. We based our model on parameters from a network meta-analysis of randomised controlled trials for IUI-OS. We compared the following three agents: CC (oral medication), Letrozole (oral medication) and gonadotrophins (subcutaneous injection).

PARTICIPANTS/MATERIALS, SETTING, METHODS: The main health outcomes were cumulative live birth and multiple pregnancy. As the procedures are identical except for the agent used, we only considered direct medical costs of the agent during four cycles. The main cost-effectiveness measures were the differences in costs divided by the differences in cumulative live birth (incremental cost-effectiveness ratio, ICER) and the probability of the highest net monetary benefit in which costs for an agent were deducted from the live births gained. The live birth rate for IUI using CC was taken from trials adhering to strict cancellation criteria included in a network meta-analysis and extrapolated to four cycles. We took the relative risks for the live birth rate after Letrozole and gonadotrophins versus CC from that same network meta-analysis to estimate the remaining absolute live birth rates. The uncertainty around live birth rates, relative effectiveness and costs was assessed by probabilistic sensitivity analysis in which we drew values from distributions and repeated this procedure 20 000 times. In addition, we changed model assumptions to assess their influence on our results.

MAIN RESULTS AND THE ROLE OF CHANCE: The agent with the lowest cumulative live birth rate over 4 IUI-OS cycles conducted within one year was CC (29.4%), followed by Letrozole (32.0%) and gonadotrophins (34.5%). The average costs per four cycles were €362, €434 and €1809, respectively. The ICER of Letrozole versus CC was €2809 per additional live birth, whereas the ICER of gonadotrophins versus Letrozole was €53 831 per additional live birth. When we assume a live birth is valued at €3000 or less, CC had the highest probability of maximally 65% to achieve the highest net benefit. Between €3000 and €55 000, Letrozole had the highest probability of maximally 62% to achieve the highest net benefit. Assuming a monetary value of €55 000 or more, gonadotrophins had the highest probability of maximally 56% to achieve the highest net benefit.

LIMITATIONS, REASONS FOR CAUTION: Our model focused on population level and was thus based on average costs for the average number of four cycles conducted. We also based the model on a number of key assumptions. We changed model assumptions to assess the influence of these assumptions on our results.

WIDER IMPLICATIONS OF THE FINDINGS: The high uncertainty surrounding our results indicate that more research is necessary on the relative effectiveness of using CC, Letrozole or gonadotrophins for IUI-OS in terms of the cumulative live birth rate. We suggest that in the meantime, CC or Letrozole are the preferred choice of agent.

STUDY FUNDING/COMPETING INTEREST(S): This work was supported by ZonMw Doelmatigheidsonderzoek, grant 80-85200-98-91072. The funder had no role in the design, conduct or reporting of this work. BWM is supported by a NHMRC Practitioner Fellowship (GNT1082548). B.W.M. reports consultancy for ObsEva, Merck KGaA and Guerbet and travel and research support from ObsEva, Merck and Guerbet. All other authors have no conflicts of interest to declare.

TRIAL REGISTRATION NUMBER: N/A.

Key words: medically assisted reproduction / unexplained subfertility / IUI / cost-effectiveness / decision analytic model / decision tree / ovarian stimulation / clomiphene citrate / Letrozole / gonadotrophins

Introduction

IUI with ovarian stimulation (IUI-OS) is commonly used to treat couples with unexplained subfertility, in particular as first-line treatment (Calhaz-Jorge *et al.*, 2017). The ideal protocols for conducting IUI-OS remain an ongoing debate among clinicians with differences between countries, clinics and settings, despite efforts for global recommendations in international guidelines (Cohlen *et al.*, 2018). A main question that remains is: which agent can best be used for OS when conducting IUI? Here, 'best' can be defined in three ways: the agent that yields the highest cumulative live birth rate, the agent that is the safest in terms of the lowest chance of a multiple pregnancy or the agent that is the most cost-effective. The current international guideline specifically identified the need for more research on the cost-effectiveness of agents used for mild stimulation in IUI protocols (Cohlen *et al.*, 2018).

Effectiveness in terms of live birth and safety in terms of multiple pregnancy have been addressed in a network meta-analysis (NMA) that pooled all available data from randomised controlled trials (RCTs) on IUI and IUI-OS (Wang *et al.*, 2020). This NMA distinguished between trials with and without strict cancellation criteria, defined as cycle cancellation when exceeding three follicles of ≥ 14 millimetre. These criteria are an effort to reduce the probability of a multiple pregnancy. In the subgroup of trials adhering to strict cancellation criteria there was no longer a statistically significant effect of gonadotrophins versus CC, in contrast to the primary analysis (Wang *et al.*, 2020). This was mostly due to the recent SUPER trial which did not find a large difference between these agents in terms of the cumulative live birth rate (Danhof *et al.*, 2018). The NMA was inconclusive as to which agent significantly increased the chance of a multiple pregnancy.

Given the current body of evidence and contemporary recommended protocol adhering to strict cancellation criteria, we are uncertain which agent to use for OS, as relative effectiveness, safety and cost-effectiveness are all unknown. A cost-effectiveness study compared gonadotrophins to CC in the SUPER trial and found that gonadotrophins were more effective but also more costly with an incremental cost-effectiveness ratio (ICER) of €17 044 (95% CI: €8998–€25 090) (Danhof *et al.*, 2019). However, this was based on a single Dutch study in which the authors did not compare CC and gonadotrophins to the commonly used alternative Letrozole, since Letrozole is unregistered in the Netherlands for this indication. More

evidence is necessary on this topic to decide on sustainable, evidence-based treatment protocols that best utilise public funds for IUI-OS with the indication of unexplained subfertility.

The aim of the present study is to assess the cost-effectiveness of stimulation agents for IUI-OS in couples with unexplained subfertility by using the most up-to-date body of evidence.

Materials and methods

To answer the research question we did not use patient data, but instead combined evidence from contemporary sources to develop a statistical model commonly referred to as a decision analytic model (Briggs *et al.*, 2006). We chose a basic population level (average) model based on probabilities known as a decision tree.

We reported the study according to the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist. We published the Excel sheet including all parameters that were used (Supplementary Data File S1) and R code used to run (Supplementary Data File S2) and analyse (Supplementary Data File S3) the model alongside this manuscript such that other researchers can readily implement the model in their own country or setting where parameters such as costs might differ.

Population and study design

We defined subfertility as not conceiving within one year of trying. Unexplained subfertility was defined as couples with no major causes explaining the fertility problem, regular menstrual cycles of length between 23 and 35 days, at least one patent fallopian tube and semen analyses with a total motile sperm count of $>10^6$ (van Eekelen *et al.*, 2017). Note that all literature used to develop our cost-effectiveness model is specific to unexplained subfertile couples. As this cost-effectiveness analysis concerns only the agents used in IUI, we chose the health economic perspective of society. The model can be easily modified for other perspectives or other locations. Note that the treatments are identical in everything except for stimulation agent, and therefore the societal perspective coincides with the perspective of health care providers.

The three agents that we compared are all commonly used for OS when conducting IUI in couples with unexplained subfertility:

clomiphene citrate (CC), Letrozole and gonadotrophins (Wang et al., 2020). We did not set certain dosages as we are dependent on the body of evidence given dosages used in studies. We assumed that strict cancellation criteria, defined as cancellation of insemination when exceeding three follicles of ≥ 14 millimetre, were adhered to.

We chose a protocol of up to six cycles of IUI-OS to be received in adherence to the current Dutch and international national protocol for unexplained subfertility (NVOG, 2010; Cohlen et al., 2018; FMS, 2020). When protocol is six cycles, most couples actually receive 4 because couples can conceive in the first or second cycle, couples drop out of IUI, couples continue to IVF before having received six cycles, etc. (Custers et al., 2007). We therefore assumed in our model that patients received, on average, four cycles, to be completed within one year.

Primary data sources

Our outcome of interest is the cumulative probability of live birth as this is the outcome that matters most to patients and quality-adjusted life years are poorly defined for this population (Barnhart, 2014; Eijkemans et al., 2017). As the baseline rate in the model, we chose to use the per-cycle rate after CC in RCTs with a protocol up to four cycles that adhered to strict cancellation criteria from the NMA (Wang et al., 2020). This NMA only encompasses RCTs that included unexplained subfertile patients. This is the most up-to-date live birth rate available for the least expensive agent, given adherence to contemporary strict cancellation criteria protocols that minimise the chance of a multiple pregnancy. The rate can be used to calculate the cumulative live birth rate over any number of cycles desired, although this represents an extrapolation. We derived the cumulative live birth rate (CLBR) over four cycles as:

$$CLBR = 1 - (1 - rate)^4. \quad (1)$$

The average rate after IUI using CC plugged into the model was 8.35%, yielding an average cumulative live birth rate over four cycles of 29.5% following the formula in (1) (Wang et al., 2020).

For the cumulative live birth rates after four cycles of Letrozole or gonadotrophins, we used the relative risks in RCTs with a protocol up to four cycles that adhered to strict cancellation criteria (Wang et al., 2020). This approach has three major advantages: first, it is based on all evidence from RCTs on IUI for unexplained subfertile couples; second, as this evidence is from RCTs, the relative risks are marginal, i.e. on the population level which coincides with the perspective of the decision analytic model; third, these relative risks are derived from intention-to-treat analyses that incorporate deviations from protocol such as natural conception or changing to a different agent in a later IUI cycle, as that would occur in clinical practice as well. The NMA reported a relative risk of Letrozole versus CC of 1.09 (95% CI: 0.76–1.57, only evidence in RCTs without strict cancellation criteria) and of gonadotrophins versus CC of 1.20 (95% CI: 0.95–1.51) (Wang et al., 2020). We thus model that all agents would lead to an average of four cycles, assuming no difference in dropout rates between the three agents (Danhof et al., 2018 Wang et al., 2020).

Following RCTs with a protocol up to four cycles that adhered to strict cancellation criteria, we expected a rate of multiple pregnancy per woman randomised of 2.5% after CC, 3% after Letrozole and 5.3% after gonadotrophins (Wang et al., 2020). A multiple pregnancy

was considered equivalent to a singleton pregnancy, i.e. a single event in terms of ‘success’ and counting towards the live birth rate. We did not use the relative risks for multiple pregnancy from the NMA as these were very imprecise and not suited for network analysis.

Costs

Costs for procedures regarding IUI treatment were considered identical for all three agents. Cost data of medication for IUI-OS using CC, Letrozole or gonadotrophins were obtained from the Dutch Formulary on Medication (Farmacotherapeutisch Kompas, 2019). Costs data regarding delivery of a multiple pregnancy were obtained from an expert panel on cost-effectiveness from the Dutch consortium for Research in Women’s Health consisting of gynaecologists, an economist and a methodologist. This group collected the total medical costs per resource unit from two university hospitals and one general hospital. These costs were averaged and then used to determine the average costs of a multiple delivery.

Medication-only costs per cycle of IUI-OS were estimated at €0.90 for CC (generic), €1.00 for Letrozole (generic) and €262 for the most commonly used gonadotrophins in the Netherlands (Gonal-F, Merck) (Farmacotherapeutisch Kompas, 2019). Thus, after four cycles, the expected costs were €3.60 for CC, €4.00 for Letrozole and €1048 for gonadotrophins.

Multiple delivery cost data were from 2013, so we applied the inflation factor reported by the Dutch government to update costs to 2019 (CBS, 2020). The incremental cost of delivery in case of a multiple pregnancy was estimated at €14 357 and was defined as the costs of a multiple pregnancy delivery minus the cost of a singleton delivery, as the latter applies to all deliveries (Lukassen et al., 2004).

All details are reported in the Model Parameters Excel sheet (Supplementary Data File S1).

Cost-effectiveness analyses

The main outcomes from the model were the cumulative live birth rate, the average costs and the probability of multiple pregnancy. Our model is based on a hypothetical population of couples with unexplained subfertility. The uncertainty around effectiveness, costs and other parameters was assessed by probabilistic sensitivity analysis in which we drew values from distributions and repeated this procedure 20 000 times. Thus, the basic form of our model is a Monte Carlo simulation that executes a decision tree, replicated many times.

The structure of our decision tree model is as follows: first, for each simulation replication, we drew values from the appropriate distributions for the relative risks, costs and the baseline rate of live birth after CC. For each agent, we applied the relative risk to the baseline rate and calculated the cumulative live birth rate over four cycles. We added the costs of multiple pregnancy deliveries to the total costs.

All parameters used, their distribution and the source of evidence are reported in Table 1 and are reported in more detail in the Excel sheet. The model simulations were performed using R version 3.6.0 (R Core Team, 2017). All code necessary to run and analyse the model is provided online at *Human Reproduction*.

We did not apply discounting procedures for economic analyses as our decision analytic model covers only four cycles of IUI, assumed to be completed within one year (Briggs et al., 2006).

Table 1 Parameters, their distribution, assumptions, evaluations of assumptions and data source as used in the decision analytic model.

Parameter	Distribution	2.5th and 97.5th percentile	Underlying assumptions	Assumptions checked by	Source (see also Supplementary Data File S1)
Costs for clomiphene citrate (CC)	$\text{Gamma}\left(\sqrt{3.60}, \frac{1}{\sqrt{3.60}}\right)$	0.40–10.20	Costs for 4 IUI-OS cycles	Probabilistic sensitivity analysis	Dutch Formulary for Medication
Costs for Letrozole	$\text{Gamma}\left(\sqrt{4.00}, \frac{1}{\sqrt{4.00}}\right)$	0.48–11.14	Costs for 4 IUI-OS cycles	Probabilistic sensitivity analysis	Dutch Formulary for Medication
Costs for gonadotrophins	N (1048,60)	930–1166	Costs for 4 IUI-OS cycles	Probabilistic sensitivity analysis	Dutch Formulary for Medication
Costs for multiple delivery	N (14 357, 250)	13 867–14 847	Costs known and increases by general price index inflation	Probabilistic sensitivity analysis	Dutch medical centre cost data
Relative risk of Letrozole versus CC (all trials)	Lognormal(ln(1.09) 0.1862)	0.76–1.57	Known relative effect that relies on the base-line rate	Probabilistic sensitivity analysis, scenario analysis	Network meta-analysis Wang et al. (2020)
Relative risk of gonadotrophins versus CC (trials up to four cycles with strict cancellation criteria)	Lognormal(ln(1.20) 0.1172)	0.95–1.51	Known relative effect that relies on the base-line rate	Probabilistic sensitivity analysis, scenario analysis	Network meta-analysis Wang et al. (2020)
Relative risk of gonadotrophins versus CC (all trials)	Lognormal(ln(1.39) 0.1204)	1.09–1.76	Known relative effect that relies on the base-line rate	Scenario analysis	Network meta-analysis Wang et al. (2020)
Probability of multiple pregnancy after CC	Beta(18.5, 721.5)	1.5–3.7%	Multiple pregnancy rates in trials in applicable to population	Probabilistic sensitivity analysis, scenario analysis	Network meta-analysis Wang et al. (2020)
Probability of multiple pregnancy after Letrozole	Beta(8.97, 290)	1.4–5.2%	Multiple pregnancy rates in trials in applicable to population	Probabilistic sensitivity analysis, scenario analysis	Network meta-analysis Wang et al. (2020)
Probability of multiple pregnancy after gonadotrophins	Beta(51.3, 916.7)	4.0–6.8%	Multiple pregnancy rates in trials in applicable to population	Probabilistic sensitivity analysis, scenario analysis	Network meta-analysis Wang et al. (2020)
Live birth rate, per cycle	Beta(140, 1536)	7.1–9.7%	Rate can be accurately extrapolated to 4 full cycles	Probabilistic sensitivity analysis, scenario analysis	Network meta-analysis Wang et al. (2020)
Live birth rate in trials up to four cycles who adhered to strict cancellation criteria	N (0.246,0.0158)	21.5–27.7%	Protocol for these trials is representative of today's IUI-OS population for unexplained	Probabilistic sensitivity analysis, scenario analysis	Network meta-analysis Wang et al. (2020)

Measures of cost-effectiveness

The main cost-effectiveness measures were the ICERs, the percentage of simulation replications for each quadrant of the cost-effectiveness plane and the probability that an agent was most likely to yield the highest net monetary benefit.

The ICER is the increase in average cost divided by the increase in the live birth rate for a more effective, but also more expensive agent (Briggs et al., 2006). We opted for a step-wise approach, comparing ICERs for increasingly more effective but also more expensive agents (Briggs et al., 2006). We calculated the percentage of simulation replications in each quadrant of the two-dimensional cost-effectiveness plane using the same step-wise approach.

For net (monetary) benefit, we expressed the gain for each agent regarding live birth in terms of a monetary value. We achieved this by multiplying the cumulative live birth rates by a monetary value between

€1 and €100 000 and deducting the total costs of that agent (Briggs et al., 2006). The result is the monetary benefit of that agent in terms of health gained minus the costs. For each simulation replication, we determined which of the three agents yielded the highest net benefit for the range of possible monetary values, then calculated the proportion of simulation replications in which each agent was the best. The results are thus interpretable as the probability that an agent yields the highest net benefit, given that we assume live birth to be valued a certain monetary value.

Scenario analyses

In addition to the primary analysis, we changed model assumptions to assess their influence on our results.

First, we used the relative risk for the live birth rate after gonadotrophins versus CC from the NMA as derived from pooled evidence on

all included RCTs instead of only RCTs with protocol up to four cycles that adhered to strict cancellation criteria. This alternative relative risk was 1.39 (95% CI: 1.09–1.76) (also mentioned in Table I).

Second, to assess the influence of the number of average cycles that couples received, we assumed a higher dropout rate leading to two cycles conducted on average instead of four, changing the average costs to €1.80 after CC, €2.00 after Letrozole and €524 after gonadotrophins and extrapolating the live birth and multiple pregnancy rate to 2 cycles. We repeated this process but now with a protocol to continue for more than six cycles, in which we assumed the average would be six cycles, changing the average costs to €5.40 after CC, €6.00 after Letrozole and €1572 after gonadotrophins and extrapolating the live birth and multiple pregnancy rate to six cycles.

Third, we assumed that the baseline live birth rate was the cumulative live birth rate over the total follow up in RCTs with protocol up to four cycles that adhered to strict cancellation criteria, which was 24.6% (Wang et al., 2020).

Fourth, we chose parameters based on a previous analysis of the individual participant data (IPD) from authors of trials that were included in the NMA and whom provided their data (Wang et al., 2020). This changed the cumulative live birth rate after CC to 29.8%, the relative risk of Letrozole versus CC to 0.80 (95% CI: 0.59–1.10) and the relative risk of gonadotrophins versus CC to 1.28 (95% CI: 1.04–1.66).

Results

The main results are displayed in Table II. The agent with the lowest cumulative live birth rate was CC (29.4%), then Letrozole (32.0%) and gonadotrophins with the highest (34.5%). The average costs were €362, €434 and €1809, respectively. This meant that Letrozole was more effective but also more costly than CC and gonadotrophins were more effective and more costly than both CC and Letrozole. The ICER for Letrozole was €2809 compared to CC. The ICER for gonadotrophins was €53 831 compared to Letrozole. The range of ICER values between the lowest 2.5% and highest 97.5% of simulation replications was extremely broad, reflecting the very high uncertainties in the relative risk estimates in addition to uncertainty in the average costs.

The percentages of simulation replications for each quadrant of the cost-effectiveness plane are presented in Table III. Despite the point estimates showing that Letrozole was more effective but also more

costly than CC, this only occurred in 44.0% of replications (north-east quadrant). In 23.3% of replications, Letrozole dominated CC, i.e. being more effective and less costly. In 21.6% of replications, Letrozole was inferior to CC, being less effective and more costly. Similarly, despite the point estimates showing that gonadotrophins were more effective but also more costly than Letrozole, this only occurred in 67.4% of replications (north-east quadrant). In the remaining 32.6%, gonadotrophins were inferior to Letrozole. These results reaffirm the finding in ICERs that the uncertainty around parameters was very high.

The net benefit curves for a range of monetary value per live birth are shown in Fig. 1.

The curve showed that when we assume a live birth to value between €1 and €3000, CC had the highest probability of maximally 65% to achieve the highest net benefit. Between €3000 and €55 000, Letrozole had the highest probability of maximally 62% to achieve the highest net benefit. Assuming a monetary value of €55 000 or more, gonadotrophins had the highest probability of maximally 56% to achieve the highest net benefit. The curve in Fig. 2 is identical to Fig. 1 but shows only the best performing agent over the range of monetary values. The probabilities to yield the highest benefit never exceeded 65%, reflecting the high uncertainty in the parameters and indicating a low certainty that the agents are truly better at these different monetary values.

In short, when assuming a live birth to be 'valued' a low, intermediate or high amount of money, the most cost-effective agent was CC, Letrozole or gonadotrophins, respectively. Unfortunately, the uncertainty surrounding results was high, not allowing for strict cut-offs for these values.

Scenario analyses

The ICER for gonadotrophins versus Letrozole was lower in the scenario analysis in which we used the higher relative risk for gonadotrophins based on all included trials in the NMA, at €19 500 (Supplementary Table S1). However, the 95% CI limits remained extremely broad. From a monetary value of approximately €20 000 onwards, gonadotrophins now were the most likely to yield the highest net benefit, with a maximum probability of approximately 81% (Supplementary Fig. S1).

The ICER for gonadotrophins versus Letrozole with 2 cycles instead of 4 was slightly lower than the primary analysis at €45 000

Table II Main outcomes from the cost-effectiveness analysis (all agents).

Agent	Cumulative probability of live birth over four cycles %	Average costs (€)	Probability of multiple pregnancy %	Incremental cost-effectiveness ratio (€) (ICER)*
Clomiphene citrate	29.4	362	2.5	0 (reference)
Letrozole	32.0	434	3.0	2809 (95% CI: -46 462 to 41 136)
Gonadotrophins	34.5	1809	5.3	53 831 (95% CI: -294 542 to 345 764)

All costs in €EUR.

*Step-wise approach first comparing Letrozole to clomiphene citrate and second gonadotrophins to Letrozole.

Table III Quadrants of the cost-effectiveness plane.

Agent comparison	North-east (more effective but more expensive)	South-west (less effective but less expensive)	North-west (less effective and more expensive i.e. inferior)	South-east (more effective and less expensive i.e. dominant)
Letrozole versus Clomiphene citrate	44.0%	11.1%	21.6%	23.3%
Gonadotrophins versus Letrozole	67.4%	0%	32.6%	0%

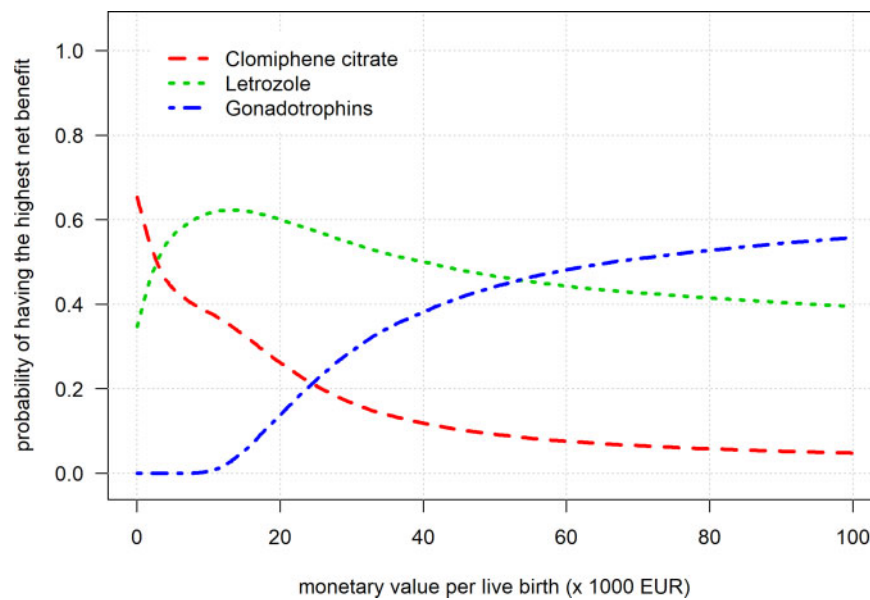


Figure 1. Net benefit curves for all agents.

(Supplementary Table SII). The ICER for six cycles instead of four was slightly higher than the primary analysis at €67 000 (Supplementary Table SIII). However, the 95% CI limits remained extremely broad. Both net benefit curves were similar to the primary analysis, with slightly shifted crossover points but identical patterns in terms of the probability that one agent yields the highest net benefit (Supplementary Fig. S2a and b).

Results for both the ICERs and the net benefit curves in the scenario analysis in which we used the cumulative live birth rate from trials that protocolled up to four cycles with adherence to strict cancellation criteria were very similar to the primary analysis (Supplementary Table SIV and Fig. S3).

CC dominated Letrozole in the scenario analysis when using parameters based on the IPD, as CC was more effective and less expensive. Gonadotrophins were significantly more effective than CC but also significantly more expensive, with an ICER of approximately €20 500 (Supplementary Table SV and Fig. S4). The net benefit curve now shows Letrozole as the worst option at any monetary value. If a live birth is assumed to be valued €15 000 or less, CC was the most likely to yield the highest net benefit, with a probability of 80% or higher. If

a live birth is assumed to be valued €33 000 or more, gonadotrophins were the most likely to yield the highest net benefit, with a probability of 80% or higher. Between €15 000 and €33 000, it was uncertain which agent yielded the highest net benefit.

Discussion

We found that if stakeholders or society as a whole consider a live birth to be valued a relatively low monetary value (€3000 or less), the least expensive but most cost-effective agent in an IUI-OS protocol in terms of net benefit was CC. If a live birth is considered to be valued between €3000 and €55 000, Letrozole was the most cost-effective agent. If a live birth is considered to be valued €55 000 or more, gonadotrophins were the most cost-effective agent. The probability that one agent was the most cost-effective never exceeded 65%, indicating the high uncertainty in model parameters and results. When we used the parameters from the IPDMA, which we consider as a selection of studies that were generally of higher quality, the relative effects were more precise and Letrozole was now found less effective than CC and

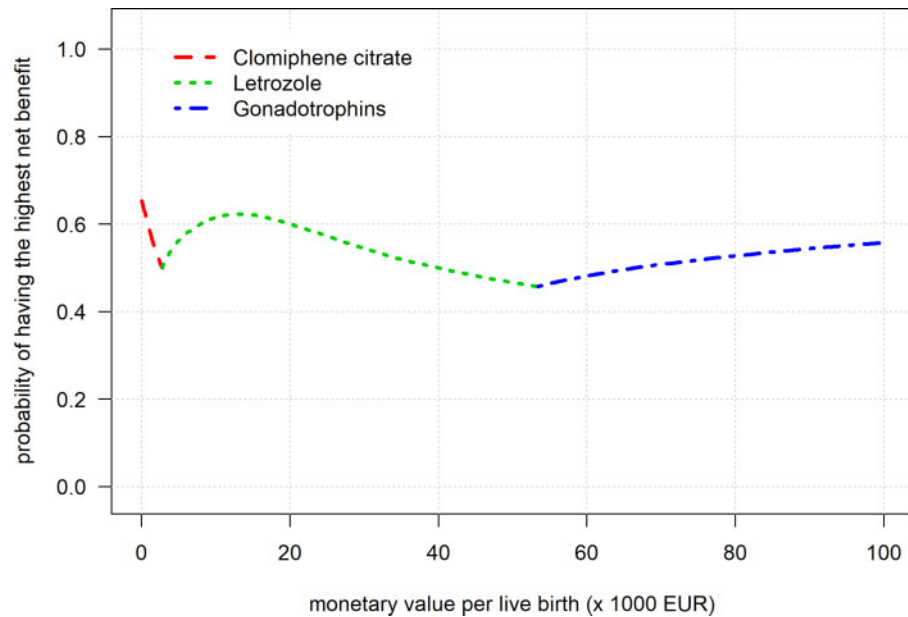


Figure 2. The cost-effective frontier for the net benefit curves.

gonadotrophins significantly more effective than CC. The net benefit curve showed CC as the best option below a monetary value of €15 000 per live birth and gonadotrophins as the best option above a monetary value of €33 000.

Strengths of this study include the use of evidence from a contemporary NMA that combined all evidence from RCTs conducted in IUI-OS for unexplained subfertility for all three relevant, often-used agents. In addition, uncertainty was incorporated by alternative assumptions in our scenario analyses as well as using distributions in a probabilistic sensitivity analysis in which the model simulation was repeated 20 000 times.

As mentioned, there was a very high uncertainty surrounding estimates for the decision analytic model shown in both the 95% CI limits for the ICER and, more informatively, in the net benefit curves. The uncertainty was highest surrounding the relative effectiveness of gonadotrophins and Letrozole as compared to CC. In the subgroup analysis for trials that protocolled up to four cycles of IUI and adhered to strict cancellation criteria, there were no direct comparisons between Letrozole and CC and thus we had to use the relative risk from the remaining trials.

Our model, although based on contemporary evidence, was not free of assumptions and was based on average costs for an average number of cycles, assumed to be completed within 1 year. This reflects the societal perspective on population level, i.e. average costs.

Most of our scenario analyses showed similar results as in the primary analysis except when we used the relative risk parameters from the IPD. This was mainly due to the inclusion of trials that did not adhere to strict cancellation criteria: Letrozole then led to fewer live births and the relative effect of gonadotrophins on live birth was further away from 1 and statistically significant, both compared to CC. There remained great uncertainty surrounding the ICER and in the net benefit curve, showing that between a monetary value of €15 000 and

€33 000 per live birth, it remained unknown which option was likely to yield the highest net benefit.

There is no consensus on what society deems “cost-effective” or “worth investing” regarding live birth, which complicates decision making. Especially in the framework of net benefit, in which a live birth on itself is declared to be of certain monetary value rather than the investment required to gain one additional live birth for option A compared to option B, this can be a sensitive topic. Preferences vary for stakeholders and perspective and, in particular, depend on country and health care system. The net benefit curve allows readers to interpret from their own perspective.

Our results beg once more for a re-evaluation of the recommended protocols for IUI-OS. In particular, our results call for more research on the relative effectiveness of both Letrozole and gonadotrophins versus CC: Letrozole as the least data was available and its effectiveness remains very uncertain and gonadotrophins as they are 291 times more expensive as a dosage per cycle as compared to CC at approximately €262 versus €0.90. Our study shows that this steep difference in price is unwarranted in terms of the probability of live birth. Couples who pay for their own treatment and medication can reduce spending considerably without evidence of decreasing their probability of live birth. Insurance companies and governments can cut costs without evidence of harming patients.

We suggest that with the current body of evidence, CC or Letrozole are the preferred choice of agent.

The results of our study can be used in discussions between clinicians, couples and policy makers to decide on a sustainable treatment protocol on IUI-OS for couples with unexplained subfertility. Our model used contemporary evidence and all necessary code is provided such that it is easy to apply in other countries where, for instance, costs might differ.

Conclusion

The high uncertainty surrounding our results indicate that more research is necessary on the relative effectiveness of using CC, Letrozole or gonadotrophins for IUI-OS in terms of the cumulative live birth rate. We suggest that in the meantime, CC or Letrozole are the preferred choice of agent.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Data availability

No new data were generated or analysed in support of this research. All code and parameters used to run and analyse the cost-effectiveness model are available at *Human Reproduction* online.

Acknowledgements

The authors wish to thank M.P. Diamond, R.S. Legro, K. Peeraer, M. Erdem, T. Dankert, and R. Ecochard, on behalf of the IUI IPDMA collaboration, for providing data from their RCTs for the IPDMA of which the results are used in a sensitivity analysis.

Authors' roles

M.v.W., M.M., F.M. and B.W.M. conceived the study. R.v.E. and M.v.W. designed the statistical analysis plan. R.W., N.A. and B.W.M. conducted the network meta-analysis and selected relevant parameters for the decision analytic model. R.v.E. programmed the decision analytic model, ran simulations and analysed the results. R.v.E. and M.v.W. drafted the manuscript. All authors contributed critical revision to the paper and approved the final manuscript.

Funding

This work was supported by ZonMw Doelmatigheidsonderzoek, grant 80-85200-98-91072. The funder had no role in the design, conduct or reporting of this work. B.W.M. is supported by a NHMRC Practitioner Fellowship (GNT1082548).

Conflict of interest

B.W.M. reports consultancy for ObsEva, Merck KGaA and Guerbet and travel and research support from ObsEva, Merck and Guerbet. No other authors report a conflict of interest.

References

Bamhart KT. Live birth is the correct outcome for clinical trials evaluating therapy for the infertile couple. *Fertil Steril* 2014;**101**:1205–1208.
 Briggs A, Claxton K, Sculpher M. *Decision Modelling for Health Economic Evaluation*. Oxford: Oxford University Press, 2006.

Calhaz-Jorge C, De Geyter C, Kupka MS, de Mouzon J, Erb K, Mocanu E, Motrenko T, Scaravelli G, Wyns C, Goossens V. Assisted reproductive technology in Europe, 2013: results generated from European registers by ESHRE. *Hum Reprod* 2017;**32**:1957–1973.
 CBS (Dutch Central Bureau for Statistics). Annual mutation of price index; from 1963. 2020. <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70936ned/table?ts=1580740889682> (10 December 2020, date last accessed).
 Cohlen B, Bijkerk A, Van der Poel S, Ombelet W. IUI: review and systematic assessment of the evidence that supports global recommendations. *Hum Reprod Update* 2018;**24**:300–319.
 Custers IM, Steures P, van der Steeg JW, van Dessel TJ, Bernardus RE, Bourdrez P, Koks CA, Riedijk WJ, Burggraaff JM, van der Veen F *et al*. External validation of a prediction model for an ongoing pregnancy after intrauterine insemination. *Fertil Steril* 2007;**88**:425–431.
 Danhof NA, van Wely M, Repping S, Koks C, Verhoeve HR, de Bruin JP, Verberg MFG, van Hooff MHA, Cohlen BJ, van Heteren CF, SUPER study group *et al*. Follicle stimulating hormone versus clomiphene citrate in intrauterine insemination for unexplained subfertility: a randomized controlled trial. *Hum Reprod* 2018;**33**:1866–1874.
 Danhof NA, van Wely M, Repping S, van der Ham Dp Klijn N, Janssen I, Rijn-van Weert JM, Twisk M, Traas MAF, Pelinck MJ *et al*. Gonadotrophins or clomiphene citrate in couples with unexplained infertility undergoing intrauterine insemination: a cost-effectiveness analysis. *Reprod Biomed Online* 2019;**40**:99–104.
 Danhof NA, Wang R, van Wely M, van der Veen F, Mol BWJ, Mochtar MH. IUI for unexplained infertility—a network meta-analysis. *Hum Reprod Update* 2020;**26**:1–15.
 Eijkemans MJC, Kersten FAM, Lintsen AME, Hunault CC, Bouwmans CAM, Roijen LH, Habbema JDF, Braat DDM. Cost-effectiveness of 'immediate IVF' versus 'delayed IVF': a prospective study. *Hum Reprod* 2017;**32**:999–1008.
 Federation of Medical Specialists in the Netherlands (FMS). Guideline unexplained subfertility 2020 (in Dutch). 2020. <https://tinyurl.com/yys6rjta> (10 December 2020, date last accessed).
 Gonal-F (Farmacotherapeutisch Kompas, Dutch Formulary on Medication. Data accessed for clomiphene, Letrozole). 2019. <https://www.farmacotherapeutischkompas.nl/bladeren/preparaat/teksten/f/follitropine> (10 December 2020, date last accessed).
 Lukassen HG, Schonbeck Y, Adang EM, Braat DD, Ziehluis GA, Kremer JA. Cost analysis of singleton versus twin pregnancies after in vitro fertilization. *Fertil Steril* 2004;**81**:1240–1246.
 NVOG. Dutch Society for Obstetrics and Gynaecology. Guideline on: subfertility. 2010. <http://bit.ly/1UhuYmV> (10 December 2020, date last accessed).
 R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2017. <http://www.R-project.org/>.
 van Eekelen R, Scholten I, Tjon-Kon-Fat RI, van der Steeg JW, Steures P, Hompes P, van Wely M, van der Veen F, Mol BW, Eijkemans MJ *et al*. Natural conception: repeated predictions over time. *Hum Reprod* 2017;**32**:346–353.
 Wang R, Danhof NA, Diamond MP, Legro RS, Peeraer K, Erdem M, Dankert T, Ecochard R, Repping S, Van der Veen F *et al*. Ovarian stimulation strategies in intrauterine insemination for unexplained or mild male factor infertility – an individual participant data meta-analysis. *Hum Reprod* 2020;**32**(Suppl 1):i132.