
Downloaded from: http://sure.sunderland.ac.uk/id/eprint/13086/

Usage guidelines
Please refer to the usage guidelines at http://sure.sunderland.ac.uk/policies.html or alternatively contact sure@sunderland.ac.uk.
Multiple pregnancy rate in patients undergoing treatment with clomifene citrate for WHO group II ovulatory disorders: a systematic review

Heather Garthwaite\textsuperscript{a,*} Jane Stewart\textsuperscript{b} and Scott Wilkes\textsuperscript{a}

\textsuperscript{a}Faculty of Health Sciences and Wellbeing, University of Sunderland, Sunderland, UK;  
\textsuperscript{b}Newcastle Fertility Centre at Life, Newcastle Hospitals NHS Foundation Trust, Newcastle-upon-Tyne, UK

CONTACT: Heather Garthwaite, Faculty of Health Sciences and Wellbeing, University of Sunderland, Sunderland, UK. Telephone: +44 (0)7896 017269; Email: heather.garthwaite@doctors.org.uk; https://orcid.org/0000-0003-2758-7967
ABSTRACT
Clomifene citrate has long been associated with a multiple pregnancy rate of up to 8-10%. Studies from which these figures originated have largely included patients who received clomifene outside of current national and international treatment guidelines. In January 2020, a systematic review of MEDLINE and relevant reference lists was conducted. Studies were included that reported multiple pregnancy rate in a cohort of patients undergoing treatment with single agent clomifene citrate for WHO group II dysovulatory subfertility. Twelve studies were included in the analysis, with a total sample size of 1,387 patients. The overall multiple pregnancy rate was 3.8% (3.6% twins/0.2% triplets); this dropped to 2.4% (all twins) when studies with a mean BMI $\geq$35kg/m$^2$ were excluded. This review suggests that contemporary single agent clomifene use following current guidelines in terms of indication, results in a lower than expected multiple pregnancy rate. Further evidence from clinical practice is required to ensure that patients are adequately informed.

KEYWORDS: clomifene, clomiphene, Clomid, multiple pregnancy, ovulation induction, PCOS
Introduction

Clomifene citrate has been the first-line medical treatment for polycystic ovary syndrome (PCOS)-related (WHO group II) subfertility for over 30 years. It acts by up-regulating follicle-stimulating hormone (FSH), which encourages follicular development (Wallach & Adashi, 1984). Potential multifollicular development raises concerns about an increased risk of multiple pregnancy, which is commonly quoted in clinical practice as 8-10% (Homburg, 2005; Macgregor et al., 1968).

Many studies reporting clomifene-associated multiple pregnancy rates have included patients treated with adjunct medications or procedures, or receiving treatment for ovarian stimulation in unexplained subfertility, rather than ovulation induction in PCOS (Dehbashi et al., 2009; Dickey & Holtkamp, 1996; Kousta et al., 1997; Macgregor et al., 1968).

The authors’ experience suggests that the observed rate of clomifene-associated multiple pregnancy is lower than commonly quoted, and that the risk presented to patients may be exaggerated. There is a lack of published work that summarises the scientific and clinical evidence available, and individual studies are often underpowered.

This systematic review is therefore designed to specifically include patients whose treatment with clomifene falls within current NICE guidance, in terms of its use for ovulation induction in WHO group II dysovulatory disorders (NICE, 2013).

Materials and methods

Design

A systematic review was conducted according to PRISMA guidelines, of primary studies that reported multiple pregnancy rate in patients undergoing treatment with clomifene for WHO group II dysovulatory subfertility.
**Eligibility criteria**

Randomised controlled trials (RCTs) and observational analytic studies presenting summary statistics were included. Only studies published in the English language and reporting raw data, including clinical pregnancy and multiple pregnancy rates, were eligible. Studies were excluded if clomifene was used for ovarian stimulation or for unexplained subfertility; if adjuncts (e.g., FSH), triggers (e.g., hCG), or intrauterine insemination were also used; or, if participants had previously been diagnosed with clomifene resistance or clomifene failure.

**Search strategy**

The search was performed on MEDLINE via PubMed with the following keywords and MeSH terms: ((clomifene OR clomiphene) AND metformin); ((clomifene OR clomiphene) AND letrozole); ((clomifene OR clomiphene) AND laparoscopic ovarian diathermy), and; ((clomifene OR clomiphene) AND predict*). Studies were also identified from the reference lists of relevant systematic reviews and published guidelines (Brown & Farquhar, 2016; NICE, 2013; Teede et al., 2018). No date range was applied. The search was completed in January 2020.

**Data extraction**

The data extracted independently from each study included author, year, and country of publication, study type, sample size, baseline characteristics, previous treatments, duration and dosing of clomifene, and pregnancy outcomes including clinical pregnancy rate, and singleton, twin and triplet pregnancy rates. Clinical pregnancy was defined as a viable pregnancy that had been diagnosed on ultrasound. Those pregnancies lost to follow-up or that miscarried prior to scanning were not included in the clinical pregnancy rate.
Analysis

Intention-to-treat analysis, based on the initial allocation of participants to a treatment arm, was conducted to avoid making any assumptions about the data. Weighting for study sample size was achieved in a pooled analysis to create an overall multiple pregnancy rate, which was displayed as a percentage. None of the included studies were examined for meta-analysis due to specific focus of this review on multiple pregnancy rate with clomifene alone, rather than an effect size comparing to an alternative intervention. The application of strict inclusion criteria according to clinical guidelines served to limit clinical heterogeneity between studies, and outcomes were not weighted.

Risk of bias

Risk of bias in reporting of multiple pregnancy rate was assessed at the study level using the Cochrane risk-of-bias (RoB 2) tool for randomised trials, or the Cochrane tool to assess risk of bias in cohort studies, as appropriate (Higgins et al., 2019). For all studies, risk-of-bias assessment focused on the effect of assignment, rather than adherence, to interventions.

Patient and public involvement

A patient advisory panel is involved in guiding the larger programme of research and helped to refine the research question for this systematic review.

Results

The initial search identified 974 non-duplicate studies, of which 828 were excluded based on the title and/or the abstract (Figure. 1). The remaining publications were downloaded and stored in a database in Endnote X8.2, and the full-text articles (n=146) were individually examined by a single reviewer according to the inclusion and exclusion criteria. Twelve
papers remained and were included in the overall analysis (Amer et al., 2017; Ayaz et al., 2013; Boostanfar et al., 2001; Eijkemans et al., 2003; Johnson et al., 2010; Legro et al., 2007; Legro et al., 2014; Nahuis et al., 2016; Palomba et al., 2005; Palomba et al., 2007; Yilmaz et al., 2006; Zain et al., 2009).

**Included studies**

The twelve studies included in this review are shown in Table 1. The studies were published between 2001 and 2017 and represent data from eight countries in Europe (n=5), America (n=3), Asia (n=3) and Oceania (n=1). Four studies were conducted by two research teams in the USA (Legro et al., 2007; Legro et al., 2014) and Italy (Palomba et al., 2005; Palomba et al., 2007), however there was no overlap between the cohorts participating in the studies. The remaining eight studies were conducted by independent research teams.

Of the twelve studies included, there were nine RCTs and one non-randomised controlled trial. In these trials, participants who received treatment with clomifene were mostly in the control arm, representing standard practice. The interventions being studied as alternatives to clomifene were letrozole (Amer et al., 2017; Legro et al., 2014), tamoxifen (Boostanfar et al., 2001), metformin (Johnson et al., 2010; Legro et al., 2007; Palomba et al., 2005; Palomba et al., 2007; Zain et al., 2009), metformin with clomifene (Ayaz et al., 2013; Johnson et al., 2010; Legro et al., 2007; Zain et al., 2009), and hCG with clomifene (Yilmaz et al., 2006). Amer et al. (2017) conducted a cross-over trial; some of the women included in our analysis had previously received letrozole but with a washout period of at least six weeks.

The remaining two studies were prospective cohort studies. Eijkemans et al. (2003) reported on clinical experience with women presenting to their fertility unit, while Nahuis et al. (2016) conducted a multicentre study of postcoital testing and its reliability in predicting success with clomifene treatment.
Per study sample size ranged from 21 to 376. The total number of participants in all included studies was 1,387. All studies recruited patients experiencing fertility issues. Eight studies only recruited patients diagnosed with PCOS (Amer et al., 2017; Ayaz et al., 2013; Johnson et al., 2010; Legro et al., 2007; Legro et al., 2014; Palomba et al., 2005; Palomba et al., 2007; Zain et al., 2009), and four studies recruited patients diagnosed with WHO group II ovulatory disorders or normogonadotrophic anovulation (Boonstanfar et al., 2001; Eijkemans et al., 2003; Nahuis et al., 2016; Yilmaz et al., 2006).

Two studies in the review included patients with a mean BMI above 35kg/m² and commented on a lower live birth rate in this cohort (Legro et al., 2007; Legro et al., 2014). Amer et al. (2017) restricted their inclusion criteria to women with a BMI ≤35kg/m², and stratified pregnancy and live birth rates in women with BMI <30 versus those with BMI 30-35, similarly reporting better outcomes in patients with a lower BMI. Johnson et al. (2010) restricted clomifene to patients with a BMI ≤32kg/m², and Palomba et al. (2005) restricted clomifene to those with a BMI ≤30kg/m². Seven studies reported BMI in the baseline characteristics but did not state an upper limit in their inclusion criteria (Ayaz et al., 2013; Boonstanfar et al., 2001; Eijkemans et al., 2003; Nahuis et al., 2016; Palomba et al., 2007; Yilmaz et al., 2006; Zain et al., 2009); these studies all reported a mean BMI below 32kg/m². Standard deviation or interquartile range for each of these is displayed in Table 1.

Participants in all studies received clomifene at a dose between 50mg and 250mg, for five days in the follicular phase of the menstrual cycle. Boonstanfar et al. (2001) did not report how many cycles of treatment their participants received, but overall, reported on 91 cycles in 40 patients. The maximum number of cycles given in other included studies ranged from one to seven.

Of the twelve included studies, eight stated that they used transvaginal ultrasound for follicular tracking during at least the first cycle of treatment with clomifene (Amer et al.,
One case was reported where a couple were advised against intercourse due to risk of multiple pregnancy, although it is unclear which treatment arm they were in (Johnson et al., 2010).

In all twelve studies, clinical pregnancy was diagnosed on ultrasound, and multiple pregnancy rate was reported as a secondary outcome.

Quality of included studies
Bias arising during randomisation or due to deviations from the intended interventions was generally low, but given the lack of comparator in this review, the relevance of this is limited. The groups that received clomifene were representative of the whole study population in each trial. There was no missing outcome data in any of the studies. Where per protocol analysis was published by the authors, adequate raw data was available for conversion to intention-to-treat analysis for the benefit of this review. All studies were deemed as being at low risk of bias in measurement of the outcome and selection of the reported result, given that multiple pregnancy rate as an absolute outcome was the only measure of interest and was a key inclusion criterion for the review.

Excluded studies
After analysis of full texts, 134 studies were excluded from the review. One of the commonest reasons for excluding studies was the use of clomifene for indications other than WHO group II anovulatory subfertility, particularly for unexplained subfertility. Studies were also commonly excluded if the protocol for clomifene administration was not clear, if clomifene was used in conjunction with another treatment, if key outcome data were missing, or if the data was a duplication.
Main outcomes

The cumulative clinical pregnancy rate across all studies was 30.1% \((n=418)\). A total of 96.2% of pregnancies were singleton, leaving a multiple pregnancy rate of 3.8% (3.6% twins/0.2% triplets; no higher-order multiple pregnancies). None of the included studies reported multiple pregnancy rates according to BMI. When studies with a known mean or median BMI \(\geq 35\text{kg/m}^2\) were excluded, the multiple pregnancy rate was 2.4% (all twins).

Discussion

This is the first published review of multiple pregnancy rates solely in patients with WHO group II anovulation who have received clomifene citrate for ovulation induction. It suggests a multiple pregnancy rate below 4%, rather than the 8-10% that is commonly quoted. The included study populations reflect those that are diagnosed in clinical practice in line with robust, standardised WHO criteria, and treated accordingly on the basis of current NICE guidance. Although only one of the twelve included studies was UK-based, the clinical management of subfertility with clomifene varies little around the world and, given the strict inclusion criteria, this is unlikely to have compromised external validity (Teede et al., 2018).

There was no date range applied to the current systematic review; this was a strength, as it demonstrated that all relevant research has only been published after 2000. The quoted multiple pregnancy risk of 8-10% was established in clinical practice by this time and does not seem to have been reviewed in light of subsequent published, contradictory rates. Potentially, rates in study populations prior to this were exaggerated by the use of adjuncts, triggers or intrauterine insemination, or by including participants with unexplained subfertility.
MacGregor et al. (1968) reported on several studies in which women were treated with clomifene for subfertility, resulting in 1,201 singleton, and 136 multiple, live births (Macgregor et al., 1968). Their conclusion that multiple pregnancy “was increased up to tenfold when conception occurred in a cycle in which clomiphene citrate was administered”, has frequently been referenced in subsequent publications. However, the doses and protocols varied significantly between the studies that they reported, and indications for treatment were also wide-ranging, including post-contraceptive amenorrhoea, lactation-amenorrhoea syndromes and psychogenic amenorrhoea; only 20% of patients had PCOS.

Due to lack of evidence, clinical guidelines on the use of clomifene are based in part on studies in which the treatment protocols used are not compliant with current recommendations. For example, the 2018 PCOS guideline endorsed by the European Society of Human Reproduction and Embryology (ESHRE) advises that the risk of twin pregnancy with clomifene citrate is 5-7%, and that, “the risk of multiple pregnancy appears to be less with letrozole, compared with clomiphene citrate” (Teede et al., 2018). Of the studies that they reference comparing clomifene with letrozole, only one could be included in the current analysis (Legro et al., 2014). The remaining seven used an hCG trigger injection with clomifene.

Similarly, NICE (2013) references fourteen studies in which clomifene is in the control arm for multiple pregnancy risk; four of these are included in the current review (Johnson et al., 2010; Legro et al., 2007; Palomba et al., 2005; Zain et al., 2009). Of the remaining ten studies, six used hCG triggers, two included clomifene-resistant patients, and in two the multiple pregnancy rate was unclear.

Finally, the Cochrane Collaboration’s 2016 review of clomifene use in ovulation induction references thirteen RCTs where clomifene was the comparator and multiple pregnancy rate is reported (Brown & Farquhar, 2016). Of these, two are included in the
current review (Boostanfar et al., 2001; Yilmaz et al., 2006), five used adjuncts to clomifene, four looked at clomifene-resistant patients, one treated women who did not have WHO group II ovulatory dysfunction, and one did not report clinical pregnancy rate.

In the current review, the studies that reported the highest multiple pregnancy rates, the only triplet pregnancy, and some of the lowest pregnancy rates overall, were those with a marked selection bias towards obese women (Legro et al., 2007; Legro et al., 2014). Risk of dizygotic twinning is increased in women with a higher BMI, and the interplay of multiple gestation with obesity makes for a high-risk pregnancy (Reddy et al., 2005). Obese women also have a poorer response to ovulation induction and require higher doses (Norman et al., 2004). The inclusion of patients with grade II obesity in fertility research is not representative of clinical practice, as most fertility centres, particularly in Europe, do not provide treatment to those with a BMI over 35.

NICE (2013) currently recommends transvaginal ultrasonography during the first cycle of clomifene, to ensure that the patient is taking a dose that minimises the risk of multiple pregnancy. Eight of the twelve studies in this review indicated that they included first-cycle ultrasound, roughly mirroring UK clinical practice and its use in 61.5% of cases (Garthwaite et al., 2018). The use of ultrasound may not correlate with an overall reduction in multiple pregnancies and further research is needed.

The inclusion of studies reporting multiple pregnancy rate as a secondary outcome was a potential limitation of this review. This reduced focus on multiple pregnancy rate as an outcome meant that the impact of factors such as clomifene dose or PCOS phenotype was not reported. The outcomes of this review may also be limited by the fact that most of the sample were selected patients suitable for an RCT.

Live birth data, which may have provided a more relevant indication of clinical outcomes, was also rarely reported. However, this would have missed multiple pregnancies
where one or more fetuses had miscarried. The use of clinical pregnancy rates avoided skewing of the data by early loss of biochemical pregnancies and ensured that all pregnancies had been diagnosed sonographically as either singleton or multiple.

In conclusion, this systematic review suggests that the risk of multiple pregnancy following clomifene treatment is lower than previously expected; it may be less than 4%, and lower than 3% when adjusted for BMI. Further data on clomifene-associated multiple pregnancy rate in the clinical setting is currently being collated through a UK-based multicentre cohort study. This will allow patients to make a better-informed decision when considering the risks of treatment with clomifene citrate.

Declarations of interest

None

Funding

This work was funded by Health Education England (North East)

References


<table>
<thead>
<tr>
<th>Authors</th>
<th>Study type</th>
<th>Country</th>
<th>Sample size (CC arm), n</th>
<th>Mean or median* age, years</th>
<th>Mean or median* BMI, kg/m²</th>
<th>CC dose, mg</th>
<th>Maximum number of CC cycles given</th>
<th>Clinical pregnancy rate, % (n)</th>
<th>Singleton pregnancy rate, % (n)</th>
<th>Twin pregnancy rate, % (n)</th>
<th>Triplet pregnancy rate, % (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amer et al. (2017)</td>
<td>RCT</td>
<td>UK</td>
<td>110</td>
<td>28.2 ± 2.3**</td>
<td>27.5 (23.1-31.3)**</td>
<td>50-100</td>
<td>7</td>
<td>37.3% (41)</td>
<td>100.0% (41)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
</tr>
<tr>
<td>Ayaz et al. (2013)</td>
<td>RCT</td>
<td>Saudi Arabia</td>
<td>21</td>
<td>31.3 ± 2.9</td>
<td>Not stated</td>
<td>50-150</td>
<td>6</td>
<td>28.6% (6)</td>
<td>100.0% (6)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
</tr>
<tr>
<td>Boonstanfar et al. (2001)</td>
<td>RCT</td>
<td>USA</td>
<td>47</td>
<td>26.5 ± 4.3</td>
<td>30.2 ± 6.2</td>
<td>50-150</td>
<td>Not stated</td>
<td>12.8% (6)</td>
<td>100.0% (6)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
</tr>
<tr>
<td>Eijkemans et al. (2003)</td>
<td>Prospective cohort study</td>
<td>Netherlands</td>
<td>240</td>
<td>27.8 ± 4.3**</td>
<td>26.9 ± 6.2**</td>
<td>50-150</td>
<td>6</td>
<td>42.9% (103)</td>
<td>96.1% (99)</td>
<td>3.9% (4)</td>
<td>0.0% (0)</td>
</tr>
<tr>
<td>Johnson et al. (2010)</td>
<td>RCT</td>
<td>New Zealand</td>
<td>36</td>
<td>28.2 ± 4.0</td>
<td>26.2 ± 3.4</td>
<td>50-150</td>
<td>6</td>
<td>38.9% (14)</td>
<td>92.9% (13)</td>
<td>7.1% (1)</td>
<td>0.0% (0)</td>
</tr>
<tr>
<td>Legro et al. (2007)</td>
<td>RCT</td>
<td>USA</td>
<td>209</td>
<td>27.9 ± 4.0</td>
<td>36.0 ± 8.9</td>
<td>50-150</td>
<td>6</td>
<td>23.9% (50)</td>
<td>94.0% (47)</td>
<td>4.0% (2)</td>
<td>2.0% (1)</td>
</tr>
<tr>
<td>Legro et al. (2014)</td>
<td>RCT</td>
<td>USA</td>
<td>376</td>
<td>28.8 ± 4.0</td>
<td>35.1 ± 9.0</td>
<td>50-150</td>
<td>5</td>
<td>21.5% (81)</td>
<td>92.6% (75)</td>
<td>7.4% (6)</td>
<td>0.0% (0)</td>
</tr>
<tr>
<td>Nahuis et al. (2016)</td>
<td>Prospective cohort study</td>
<td>Netherlands</td>
<td>152</td>
<td>29.2</td>
<td>23.4</td>
<td>50-150</td>
<td>6</td>
<td>38.8% (59)</td>
<td>100.0% (59)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
</tr>
<tr>
<td>Palomba et al. (2005)</td>
<td>RCT</td>
<td>Italy</td>
<td>50</td>
<td>25.9 ± 2.7</td>
<td>26.3 ± 2.8</td>
<td>150</td>
<td>6</td>
<td>32.0% (16)</td>
<td>100.0% (16)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
</tr>
<tr>
<td>Palomba et al. (2007)</td>
<td>Non-randomised controlled trial</td>
<td>Italy</td>
<td>40</td>
<td>29.0 (20.0-41.0)</td>
<td>26.5 (15.8-34.1)</td>
<td>50-250</td>
<td>6</td>
<td>45.0% (18)</td>
<td>94.4% (17)</td>
<td>5.6% (1)</td>
<td>0.0% (0)</td>
</tr>
<tr>
<td>Yılmaz et al. (2006)</td>
<td>RCT</td>
<td>Turkey</td>
<td>65</td>
<td>26.7 ± 3.2</td>
<td>24.2 ± 2.6</td>
<td>50</td>
<td>1</td>
<td>27.7% (18)</td>
<td>94.4% (17)</td>
<td>5.6% (1)</td>
<td>0.0% (0)</td>
</tr>
<tr>
<td>Zain et al. (2009)</td>
<td>RCT</td>
<td>Malaysia</td>
<td>41</td>
<td>29.6 ± 4.4</td>
<td>32.9 ± 4.2</td>
<td>50-200</td>
<td>6</td>
<td>14.6% (6)</td>
<td>100.0% (6)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
</tr>
<tr>
<td><strong>TOTALS</strong></td>
<td></td>
<td></td>
<td>1,387</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30.1% (418)</td>
<td>96.2% (402)</td>
<td>3.6% (15)</td>
<td>0.2% (1)</td>
</tr>
</tbody>
</table>
Table 1. Summary of studies included in systematic review.

*Age and BMI expressed as mean, mean ± SD, or median (IQR), depending on how they appeared in the original source. Baseline characteristics refer to clomifene citrate arm unless double asterisked (**), in which case they refer to all patients in the study (all arms). Abbreviations: BMI – body mass index; CC – clomifene citrate; IQR – interquartile range; RCT – randomised controlled trial; SD – standard deviation
Figure 1. PRISMA flow diagram of literature search for primary studies reporting multiple pregnancy rate in a cohort undergoing treatment with single-agent clomifene citrate for WHO group II dysovulatory subfertility.

IUI – intrauterine insemination, WHO – World Health Organization