Adverse obstetric outcomes after local treatment for cervical preinvasive and early invasive disease according to cone depth: systematic review and meta-analysis

Maria Kyrgiou,1,2 Antonios Athanasiou,3 Maria Paraskevaidi,1 Anita Mitra,1,2 Ilkka Kalliala,1 Pierre Martin-Hirsch,4,5 Marc Arbyn,6 Phillip Bennett,1,2 Evangelos Paraskevaidis3

ABSTRACT

OBJECTIVE
To assess the effect of treatment for cervical intraepithelial neoplasia (CIN) on obstetric outcomes and to correlate this with cone depth and comparison group used.

DESIGN
Systematic review and meta-analysis.

DATA SOURCES
CENTRAL, Medline, Embase from 1948 to April 2016 were searched for studies assessing obstetric outcomes in women with or without previous local cervical treatment.

DATA EXTRACTION AND SYNTHESIS
Independent reviewers extracted the data and performed quality assessment using the Newcastle-Ottawa criteria. Studies were classified according to method and obstetric endpoint. Pooled risk ratios were calculated with a random effect model and inverse variance. Heterogeneity between studies was assessed with I² statistics.

MAIN OUTCOME MEASURES
Obstetric outcomes comprised preterm birth (including spontaneous and threatened), premature rupture of the membranes, chorioamnionitis, mode of delivery, length of labour, induction of delivery, oxytocin use, haemorrhage, analgesia, cervical cerclage, and cervical stenosis. Neonatal outcomes comprised low birth weight, admission to neonatal intensive care, stillbirth, APGAR scores, and perinatal mortality.

RESULTS
71 studies were included (6 338 982 participants: 65 082 treated/6 292 563 untreated). Treatment significantly increased the risk of overall (<37 weeks; 10.7% v 5.4%; relative risk 1.78, 95% confidence interval 1.60 to 1.98), severe (<32-34 weeks; 3.5% v 1.4%; 2.40, 1.92 to 2.99), and extreme (<28-30 weeks; 1.0% v 0.3%; 2.54, 1.77 to 3.63) preterm birth. Techniques removing or ablating more tissue were associated with worse outcomes. Relative risks for delivery at <37 weeks were 2.70 (2.14 to 3.40) for cold knife conisation, 2.11 (1.26 to 3.54) for laser conisation, 2.02 (1.60 to 2.55) for excision not otherwise specified, 1.56 (1.36 to 1.79) for large loop excision of the cervix (LLETZ), 1.54 (1.26 to 1.88) for repeated excision, 1.52 (1.26 to 1.84) for ablation not otherwise specified. Compared with no treatment, the risk of preterm birth was higher in women who had undergone more than one treatment (13.2% v 4.1%; 3.78, 2.65 to 5.39) and with increasing cone depth (≥10-12 mm: 7.1% v 3.4%; 1.54, 1.09 to 2.18; ≥10-12 mm: 9.8% v 3.4%, 1.93, 1.62 to 2.31; ≥15-17 mm: 10.1% v 3.4%; 2.77, 1.95 to 3.93; ≥20 mm: 10.2% v 3.4%; 4.91, 2.06 to 11.68). The choice of comparison group affected the magnitude of effect. This was higher for external comparators, followed by internal comparators, and ultimately women with disease who did not undergo treatment. In women with untreated CIN and in pregnancies before treatment, the risk of preterm birth was higher than the risk in the general population (5.9% v 5.6%; 1.24, 1.14 to 1.35). Spontaneous preterm birth, premature rupture of the membranes, chorioamnionitis, low birth weight, admission to neonatal intensive care, and perinatal mortality were also significantly increased after treatment.

CONCLUSIONS
Women with CIN have a higher baseline risk for prematurity. Excisional and ablative treatment further increases that risk. The frequency and severity of adverse sequelae increases with increasing cone depth and is higher for excision than for ablation.

Introduction
The mean age of women undergoing local treatment for cervical preinvasive cervical disease (cervical intraepithelial neoplasia or CIN) is similar to the age of women having their first child. Local cervical treatment...
has been correlated to an increased risk of preterm birth, perinatal morbidity, and mortality in a subsequent pregnancy. The underlying mechanism is unclear; hypotheses include immunomodulation relating to infection with human papillomavirus (HPV) affecting parturition pathways and acquired “mechanical weakness” secondary to loss of cervical tissue.

In England alone in 2013-14, about 3.6 million women aged 25-64 attended for cervical screening, and over 23,800 cervical procedures were carried out, nearly all in an outpatient setting. In contrast, in the United States there are about 400,000 cases of preinvasive disease a year. The regulations in colposcopy are more liberal, leading to wide variation in clinical practice. In Germany, treatment for CIN is still commonly performed with the cold knife under general analgesia. The long term sequelae of treatment therefore remain an important international issue to healthcare professionals and women, whatever the clinical setting.

Since the first systematic review almost a decade ago on reproductive risk associated with treatment more than 50 observational studies have been published confirming or disputing these associations; some of these reported data from large population based datasets. Individual attempts to synthesise parts of this rapidly evolving evidence base in small systematic reviews and meta-analyses reached contradictory conclusions and initiated debates and confusion within the scientific community. Whether these discrepancies were due to questionable quality of some of these primary and secondary studies or to differences in the explored comparisons, the subject is open to a definitive comprehensive high quality synthesis of the existing evidence that will be highly informative to women, clinicians, and policy makers.

Media publicity has heightened public awareness that treatment for cervical precancer is associated with increased reproductive morbidity. There has been a substantial increase in inquiries from patients and clinicians on the risks associated with different treatment techniques and cone depths and as to how this risk can be managed and prevented. With a rapidly evolving evidence base and lack of a robust synthesis of the published literature, these questions are becoming increasingly difficult to answer.

We carried out a systematic review and meta-analysis to explore the impact of treatment for cervical preinvasive and early invasive disease on obstetric outcomes and how this risk could be modified by the cone depth and comparison group.

Methods
Inclusion criteria and outcomes
We included all studies that reported on obstetric outcomes (over 24 weeks’ gestation) in women who had previously received local cervical treatment for CIN or early invasive cervical cancer compared with outcomes in women with no history of treatment. Studies reporting on the outcomes after two or more treatments were also included. The interventions included any type of treatment: excisional (cold knife conisation; laser contraformation; needle excision of the transformation zone, also known as straight wire excision; large loop excision of the transformation zone, also known as loop electrosurgical excision procedure) or ablative (laser ablation; radical diathermy; cold coagulation; and cryotherapy). In studies that reported on the impact of several techniques, when possible we extracted data for each specific method. If the outcomes were not reported separately for each technique, we analysed the intervention under broader terms—that is, excisional treatment not otherwise specified, ablative treatment not otherwise specified, and treatment not otherwise specified.

Women were included irrespective of the grade of the lesion for both squamous and glandular intraepithelial neoplasia. We excluded studies that did not include an untreated reference population, compared different treatment techniques without an untreated control, or compared outcomes for treatments performed during pregnancy.

Studies were included irrespective of the type of untreated reference population that could have been drawn from one of the following sources: external group from general population that was mostly matched or adjusted for confounders; internal group with self matching of the pregnancies for the same women before and after treatment; internal group of women who had also delivered before treatment; women undergoing colposcopy with or without CIN/biopsy but no treatment; women with high grade disease but no treatment (high grade squamous intraepithelial lesion).

We assessed obstetric outcomes of pregnancies progressing beyond 24 weeks’ gestation. We examined both maternal and neonatal outcomes. The maternal outcomes included overall (<37 weeks’ gestation), severe (<32-34 weeks), and extreme (<28-30 weeks) prematurity (all preterm birth, iatrogenic and spontaneous). We also assessed preterm birth in singleton and multiple pregnancies, in nulliparous and parous women, for single and repeat cones, for different cone depths and volumes, and for different comparison groups. We further assessed other maternal outcomes that included: overall (<37 weeks’ gestation), severe (<32-34), and extreme (<28-30) spontaneous prematurity—that is, non-iatrogenic; threatened preterm birth; premature rupture of the membranes; chorioamnionitis; mode of delivery (caesarean section, instrumental deliveries); length of labour (precipitous, prolonged); induction of labour or use of oxytocin; haemorrhage (antepartum, postpartum); analgesia (epidural, pethidine, not otherwise specified); cervical stenosis; and cervical cerclage. The neonatal outcomes included: low birth weight (<2500 g, <2000 g, <1500 g, and <1000 g), admission to neonatal intensive unit, perinatal mortality, stillbirth, and Apgar score.

When there was heterogeneity in the cut offs used in different studies for cone depth and classification of prematurity, we grouped these together when possible (that is, 32-34 weeks included both cut offs; 10-12 mm cone depth included studies grouping at both these cut offs with and without the values equal to these numbers).
Literature search, data extraction, and risk of bias

We searched three electronic databases (CENTRAL, Medline, and Embase) and targeted reports published between 1948 and April 2016. We used keywords including “cervical intraepithelial neoplasia (CIN),” “cervical cancer,” “LEEP or LLETZ,” “conisation,” “excision,” “pregnancy,” “obstetric,” “preterm birth,” and “prematurity.” The full strategy is included in appendix 1. In an attempt to identify any articles missed by the initial search or any unpublished data, we hand searched the references of the retrieved articles and meta-analyses and the proceedings of relevant conferences. There was no language restriction.

From each study, we extracted data on the study design and setting, the study population, the interventions examined, the comparison group, the quality of the data and risk of bias, and the outcomes assessed. From each study and for each outcome we retrieved the number of events in treated and untreated women. If necessary, we contacted authors to obtain additional data if the numbers provided in the published report did not allow sufficient precision in the data extraction.

We used the Newcastle-Ottawa score to formally assess the quality of non-randomised cohort studies, according to the MOOSE checklist. This scoring system assesses the cohort selection, comparability, and assessment of outcomes to give a maximum score of 9 (highest quality).

Two investigators (MK, AA) independently performed the literature search, assessed the eligibility and quality of the retrieved papers, and performed the data extraction. The two authors then compared the results and disagreements were resolved by discussion. If required, consensus was reached with the involvement of a third investigator (MA).

Data synthesis and assessment of heterogeneity

We calculated the risk ratios and 95% confidence intervals for each reported outcome in the treated versus untreated women for dichotomous outcomes using Cochrane Revman 5 software. We used a random effect model and inverse variance weighting for all meta-analyses. In studies with multiple treatment groups, we proportionally divided the “shared” comparison group into the number of treatment groups; we treated comparisons between each treatment group and the split comparison group as independent comparisons. If a study presented data for more than one comparison group, we used the external comparison group of women with or without disease in preference to internal controls. If data were not of suitable quality for meta-analysis, we reported the results as a narrative in the text of the review.

We assessed heterogeneity between studies with the Cochran Q test, visual inspection of forest plots, estimation of the percentage of heterogeneity between studies that cannot be ascribed to sampling variation (I² statistic), and a formal test of the significance for heterogeneity. If there was evidence of substantial heterogeneity, the possible reasons for this were investigated and reported.

We performed a series of subgroup analyses. We analysed the data separately for each treatment, in groups of ablative and excisional techniques, and as a whole irrespective of the type of method used. We further analysed the data according to the cone depth. Given the non-randomised nature of the included studies, we assessed whether the choice of comparison group affected the risk estimate for each outcome and overinflated the effect of treatment that could be partly attributed to other confounders. We therefore distinguished the different untreated comparison groups used across studies and performed subgroup analyses for the risk of preterm birth for each individual comparator (external; internal (self matching); internal (pregnancies before treatment); colposcopy but no treatment; high grade squamous intraepithelial lesion but no treatment). Furthermore, we performed sensitivity analysis to assess the impact of the quality of the studies on some selected outcomes. We calculated the median score from the Newcastle-Ottawa scale and performed sensitivity analysis for studies that scored more than the median. We performed subgroup analyses based on the cohort selection in the Newcastle-Ottawa score (truly or somewhat representative) and the comparability of the groups (those with scores of 1 or 2). Finally, we performed meta-regression analysis to assess the impact of some factors on the risk of preterm birth (<37 weeks). These included the quality of the studies (based on the Newcastle-Ottawa score); year of study (1979-89, 1990-99, 2000-09, 2010-15); type of treatment (excision or ablation); type of comparator (external, internal-pregnancies before treatment, internal-self matching, CIN but no treatment, high grade squamous intraepithelial lesion but no treatment).

Patient involvement

Patients and the wider public were involved from the outset through informal interviews in the clinic and through patient advocate representative bodies. The research questions and outcomes were developed based on the patients’ concerns and priorities. Patients were not involved in the interpretation of results or writing of the article. The results will be disseminated to the lay audience through the authors’ involvement with charities and through public presentations.

Results

We identified 406 potentially eligible studies that fulfilled the inclusion criteria of this review. No unpublished studies were identified. We excluded studies without an untreated reference population, studies that included women treated during pregnancy, studies assessing fertility and early pregnancy outcomes below 24 weeks’ gestation, studies assessing outcomes after treatment in high risk populations, and studies assessing the impact of CIN on outcomes without information as to whether treatment was performed. Figure 1 shows more details of the literature search and the reasons for exclusion.

Table A in appendix 2 shows detailed characteristics of the included studies and the outcomes examined. Most studies were retrospective, with only five prospective reports. All were cohort studies, apart from one
case-control study by Castanon and colleagues. There were no randomised controlled studies. Fourteen studies examined the impact of cold knife conisation, 11 28 30 32 34 37 60-62 87 89 10 11 12 studied laser conisation, 42 46-47 49 51 52 56 76 78 one studied needle excision of the transformation zone, 13 34 studied large loop excision of the transformation zone, 21 39-41 44-45 50 55 60 62 63 65-69 73 74 76 81-86 88-90 91 eight studied laser ablation, 35 38 39 47 54 56-62 one studied radical diathermy, 62 two studied cryotherapy, 71 60 16 studied excision not otherwise specified, 5 12 14 15 53 64 70-72 75 78-79 84 90 93 five studied ablation not otherwise specified, 12 14 53 70 87 and three studied treatment not otherwise specified. 36 43 92 There were five types of untreated comparison groups. Some used an external comparator, 51 12 13-15 28 29 33 35-36 45-47 54 57-60 61 62 65-67 69-71 82 one studied cryotherapy, 31 five studied ablation not otherwise specified, 28 33 34 35-37 39-41 43 44 45-47 49 50 53 55-60 61 85 we included six studies comparing different cone volumes. 58 73 74 85 92 93 others compared with the pregnancies before treatment in the treated population (internal), 33 55 58 62-66 some compared with women who underwent colposcopy with or without CIN and/or biopsies. We excluded one study that compared women with high grade disease but no treatment. 35 36 56 62-63 65-67 88 89 91 93 and some with women with high grade disease but no treatment. 53 70 All studies that used an external comparison group either matched for known risk factors or performed regression analysis to control for known confounders; four studies did not control for any confounders.43 61 65 Table B in appendix 2 provides more details on the quality assessment for observational studies with the Newcastle-Ottawa score. Most studies scored 8 or 9 points, 10 scored 7 30 35 34 45-47 50 56 72 76 and two scored 6. 38 65

Maternal outcomes
The risk of preterm birth was significantly increased after cervical treatment (table 1). For all treatment types, this was the case for overall prematurity at less than 37 weeks’ gestation (relative risk 1.78, 95% confidence interval 1.60 to 1.98), for severe prematurity less than 32-34 weeks’ gestation (2.40, 1.92 to 2.99), and extreme prematurity less than 28-30 weeks’ gestation (2.54, 1.77 to 3.63) (table 1). Figure 2 shows the risk associated with LLETZ versus no treatment. The forest plot in appendix 3 shows the risks for all treatment techniques versus no treatment. The magnitude of the effect of treatment was higher for more radical treatment techniques and for excision rather than ablation. More specifically, the risk of preterm birth at less than 37 weeks’ gestation was higher for cold knife conisation (2.70, 2.14 to 3.40), laser conisation (2.11, 1.26 to 3.54), excision not otherwise specified (2.02, 1.60 to 2.55), large loop excision of the transformation zone (1.56, 1.36 to 1.79), and ablation not otherwise specified (1.46, 1.27 to 1.66). Similar trends were noted for severe and extreme prematurity.

Treatment was also associated with an increased risk of preterm birth for women with multiple pregnancies for some but not all treatments (table C, appendix 2). Have asked author, but the results were inconsistent because of the small number of studies. The impact of treatment was no different for nulliparous and multiparous women (data not shown). The effect of multiple as opposed to single treatments on the risk of prematurity was substantially higher in comparisons with untreated women (relative risk 3.78 (95% confidence interval 2.65 to 5.39) for repeated treatment and 1.75 (1.49 to 2.06) for single treatment, table 2). Compared with no treatment, the relative risk of preterm birth for two excisional treatments not otherwise specified was as high as 5.48 (2.68 to 11.24) and that of two loop excisions as high as 2.81 (2.33 to 3.39).

The magnitude of the effect according to the cone depth showed that the risk increases progressively with increasing cone depth (table 3; fig 3) or cone volume (table 3). The risk for treated versus untreated women was significantly higher for women with cone depth ≤ 10-12 mm (relative risk 1.54, 95% confidence interval 1.09 to 2.18). The magnitude of effect increased with increasing cone depth (1.93 (1.62 to 2.31) for ≥ 10-12 mm, 2.77 (1.95 to 3.53) for ≥ 15-17 mm, and 4.91 (2.06 to 11.68) for ≥ 20 mm; table 3). The trend was similar with increasing cone volume (2.25 (1.09 to 4.66) for ≤ 6 cc and 13.9 (5.09 to 37.98) for ≥ 6 cc; table 3). Further analyses of the individual cone depth cut offs not grouped together showed similar results (data not shown).

The comparison of women treated with different cone depths showed that deeper excisions significantly increased the risk of preterm birth compared with less deep excisions, and the magnitude of the effect increased in deeper cones. The relative risk were 1.54 (95% confidence interval 1.31 to 1.80) for ≥ 10-12 mm, 1.82 (1.47 to 2.26) for ≥ 15-17 mm, and 1.82 (1.47 to 2.26) for ≥ 20 mm, v < 20 mm (fig 4). Full data are also provided in table D, appendix 2. The findings were similar for comparison of cone volumes (2.04 (0.95 to 4.39) for ≥ 3-4 cc and 3.3-4 cc (15.0% v 7.3%, one study, 278 women); 6.18 (2.53 to 15.13) for ≥ 6 cc v ≤ 6 cc (50.0% v 8.1%, one study, 278 women).

We assessed the impact that the choice of comparison group can have on the magnitude of effect in a subgroup analysis that classified different studies according to the comparator used (table D). The results suggested that treatment significantly increased the
risk of preterm birth at less than 37 weeks’ gestation irrespective of the comparison group used. The magnitude of effect was higher when an external comparison group was used (relative risk 1.93, 95% confidence interval 1.71 to 2.17), followed by internal comparators (1.52 (1.17 to 1.97) for self matching and 1.42 (1.01 to 1.99) for pregnancies before treatment), and ultimately women with disease but no treatment (1.27, 1.14 to 1.41). In women with untreated CIN, and in pregnancies before treatment, the risk of preterm birth compared with general population was significantly higher (1.24, 1.14 to 1.35). The subgroup analysis of the risk of preterm birth according to cone depth and comparison group showed a similar direction of effect, although for cone depth ≤10 mm the difference became insignificant. The number of studies was small for many comparisons. For treated versus untreated women with CIN there were four studies for cone depth ≤10-12 mm (63165 women, 7.0% v. 5.0%, relative risk 1.11, 95% confidence interval 0.85 to 1.43), four studies for cone depth >10-12 mm (45275 women, 9.6% v. 5.0%, 1.52, 1.37 to 1.68), three studies for cone depth ≥15-17 mm (33934 women, 9.6% v. 4.3%, 2.30, 1.57 to 3.35), and two studies for cone depth ≥20 mm (32717 women, 9.3% v. 4.2%, 4.32, 0.93 to 20.03) (table E, appendix 2). Furthermore, the sensitivity analysis that excluded studies that scored below the median Newcastle-Ottawa score (8.3) did not change the results of the analysis; similarly the results did not change when we excluded studies that scored ≤7 and ≤6 (data not shown). The subgroup analyses of studies based on the cohort selection or the comparability of the comparison groups showed similar direction and magnitude of effect (data not shown). The univariate meta-regression analysis suggested that the type of treatment and comparator significantly affected the risk of preterm birth, although the type of treatment and Newcastle-Ottawa score did not. These factors remained significant in a multivariate regression analysis. When we performed further meta-regression restricting only to excisional treatments and using as a comparator women with cold-knife conisation/needle excision of transformation zone, NOS not eligible; NETZ=needle excision of transformation zone; NOS=no other specified; RD=radical diathermy.

If study had more than one comparison groups, we used external groups (external general, external untreated women who had colposcopy+/CIN+/biopsy, women with HSIL but no treatment) in preference to internal comparators (self matching or pregnancies before treatment).

In cases of heterogeneity in cut-offs used for classification of prematurity, these were grouped together when possible (for instance, 32-34 or 28-30 weeks included both cut offs).
excisional than ablative techniques and with more radical treatment, although the number of studies assessing each individual treatment method was often small. Cervical treatment was associated with an increased risk of spontaneous overall, severe, and extreme preterm birth (<37 weeks: 14 studies, 1024731 women, 7.0% v 3.7%; relative risk 1.76, 95% confidence interval 1.47 to 2.11; <32-34 weeks: seven studies, 655675 women, 1.8% v 0.6%; 2.63, 1.91 to 3.62; <28 weeks: two studies, 626670 women, 0.6% v 0.2%, 3.18, 1.64 to 6.16) and admissions for threatened preterm birth (five studies, 903 women, 9.1% v 3.2%, 2.44, 1.37 to 4.33). The risk (<37 weeks) was higher for cold knife conisation (3.5, 2.05 to 6.05) followed by large loop excision of the transformation zone (1.60, 1.22 to 2.08), and ablation not otherwise specified (1.42, 1.20 to 1.70). Needle excision of the transformation zone and laser ablation were each assessed in only one study. There was substantial heterogeneity for the comparisons assessing all gestational categories (P<0.05).

The risk of premature rupture of membranes (<37 weeks: 21 studies, 477011 women; 6.1% v 3.4%, relative risk 2.36, 95% confidence interval 1.76 to 3.17) and chorioamnionitis (four studies, 29198 women, 3.5 v 1.1%; 3.43, 1.36 to 8.64) was also increased after treatment. Risk was higher after cold knife conisation (4.11, 2.05 to 8.25) followed by large loop excision of the transformation zone (2.15, 1.48 to 3.12). Needle excision of the transformation zone was assessed in only one study, and laser ablation did not significantly affect the risk but was assessed in only two studies.

The mode of delivery (cesarean section or instrumental delivery), the length of labour (precipitous or prolonged), the use of analgesia (epidural, pethidine, or other), the rate of induction of labour (with or without oxytocin), cervical stenosis, and haemorrhage (antenatal or postpartum) were not affected by treatment. As expected, the rate of cervical cerclage insertion was higher for treated than non-treated women (eight studies, 141300 women, 4.0% v 0.7%, relative risk 14.29, 95% confidence interval 2.85 to 71.65) and more so for cold knife conisation (31.42, 2.32 to 426.2), large loop excision of the transformation zone (11.0, 0.64 to 190), or excisional treatment not otherwise specified (42.45, 28.99 to 62.16).

**Neonatal outcomes**
More than 30 studies assessed one or more neonatal outcomes (table G, appendix 2). Cervical treatment (excisional or ablative) was associated with a significant

---

### Table 1

<table>
<thead>
<tr>
<th>Study</th>
<th>No of events/total</th>
<th>LLETZ</th>
<th>Untreated</th>
<th>Risk ratio (95% CI)</th>
<th>Weight (%)</th>
<th>Risk ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gunasekera 1992</td>
<td>0/22</td>
<td>0/22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blomfield 1993</td>
<td>7/40</td>
<td>9/80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haffenden 1993</td>
<td>15/152</td>
<td>14/152</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Braet 1994</td>
<td>10/78</td>
<td>4/78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cruickshank 1995</td>
<td>14/147</td>
<td>15/295</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paraskevaidis 2002</td>
<td>11/28</td>
<td>3/28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sadler 2004</td>
<td>44/278</td>
<td>38/309</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tan 2004</td>
<td>13/119</td>
<td>11/119</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Samson 2005</td>
<td>44/558</td>
<td>14/558</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acharya 2005</td>
<td>9/79</td>
<td>17/158</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crane 2006</td>
<td>10/75</td>
<td>1/46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Himes 2007</td>
<td>11/114</td>
<td>127/962</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brunsma 2007</td>
<td>11/69</td>
<td>11/125</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noehr 2009</td>
<td>530/8180</td>
<td>14758/434520</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Werner 2010</td>
<td>35/511</td>
<td>17445/240348</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ortoli 2010</td>
<td>55/572</td>
<td>24265/9065</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andia 2011</td>
<td>19/189</td>
<td>10/189</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lima 2011</td>
<td>4/18</td>
<td>2/36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simoens 2012</td>
<td>12/52</td>
<td>6/104</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poon 2012</td>
<td>41/473</td>
<td>1156/25772</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frega 2013</td>
<td>26/406</td>
<td>19/379</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heinonen 2013</td>
<td>547/7636</td>
<td>30151/658179</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frey 2013</td>
<td>111/598</td>
<td>178/1140</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guo 2013</td>
<td>10/48</td>
<td>8/39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martyn 2015</td>
<td>20/278</td>
<td>6/191</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stout 2015</td>
<td>115/598</td>
<td>178/1129</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total (95% CI) 1724/21318 66607/1424023

Test for heterogeneity: $\chi^2=0.05, \chi^2=78.52, df=26, P<0.001, I^2=69%$

Test for overall effect: $z=6.38, P<0.001$

---

**Fig 2 | Meta-analysis of studies on preterm birth (<37 weeks) in women treated with large loop excision of transformation zone versus untreated women**
increase in adverse neonatal outcomes compared with outcomes in women who did not undergo treatment (comparison group not specified). The association with adverse neonatal events was stronger and more common for excisional rather than ablative techniques and with increasing treatment radicality, although the number of studies for each individual treatment technique was often limited.

More specifically, cervical treatment overall was associated with an increased risk of low birth weight (<2500 g: 30 studies, 1348206 women, 7.9% v 3.7%, relative risk 1.81, 95% confidence interval 1.58 to 2.07; <1500 g: five studies, 76836 women, 2.0% v 0.9%, 3.00, 1.54 to 5.85), admission to a neonatal intensive unit (eight studies, 2557 women, 12.6% v 8.9%, 1.45, 1.16 to 1.81), and perinatal mortality (23 studies, 1659433 women, 0.9% v 0.7%, 1.51, 1.13 to 2.03). There was significant heterogeneity between studies for perinatal mortality (P = 0.04, I² = 36%).

The rate of neonates with birth weight <2500 g was significantly higher for women treated with cold knife conisation (five studies, 30304, relative risk 2.51, 95% confidence interval 1.78 to 3.53), large loop excision of the transformation zone (12 studies, 3357, 2.11, 1.51 to 2.94), excisional (10 studies, 823648, 2.01, 1.62 to 2.49) or ablative (four studies, 483402, 1.36, 1.19 to 1.55) treatment not otherwise specified but not so for laser ablation (1.07, 0.59 to 1.92), although for that comparison there were only four studies with a total of 1104 participants. The rate of admission to neonatal intensive care was assessed only for excisional techniques and was significantly increased after large loop excision of the transformation zone (five studies, 1994 women, 1.42, 1.01 to 1.99). Perinatal mortality was significantly increased overall and for excisional technique not otherwise specified (five studies, 820028, 1.85, 1.02 to 3.36) but not for the individual techniques, possibly because of the limited number of studies and the low prevalence of the outcome. Subgroup analysis according to the different comparison groups or cone depths was not possible because of the limited number of studies assessing each outcome.

**Discussion**

**Main findings**

The knowledge that local treatment for cervical precancer, particularly excisional, increases the risk of preterm birth has led to major changes in clinical practice. With a rapidly evolving evidence base and inconsistencies in the published literature,34 15 17 18 66 113 a high quality synthesis of the evidence should be available for effective counselling of patients at colposcopy and antenatal clinics.

This meta-analysis shows that any local cervical treatment for preinvasive or early invasive disease increases the risk of preterm birth and adverse sequelae in a subsequent pregnancy, although the impact of small excisions, as opposed to just having the disease, remains uncertain and is likely to be small. Cervical treatment was found to be associated with an increased risk of overall, severe, and extreme prematurity, spontaneous preterm birth, threatened preterm labour, premature rupture of the membranes, chorioamnionitis, low birth weight, neonatal admission, and perinatal death. The rate of cervical cerclage was unsurprisingly substantially increased in treated women compared with untreated controls. Treatment equally affected outcomes for nulliparous as well as parous women and singleton and multiple pregnancies. The mode of delivery, length of labour, induction rate, use of analgesia, rate of stenosis, and haemorrhage were not significantly affected.
The magnitude of the effect of treatment was higher for more radical techniques (such as cold knife conisation followed by large loop excision of the transformation zone and laser ablation) and for excision rather than ablation. Multiple conisations increased the risk of preterm birth fourfold compared with untreated controls overall. Subgroup analyses clearly showed that the risk of preterm birth is directly correlated with the cone dimensions (depth/volume) and progressively increases with increasing cone depth (“dose effect”). Although the risk was increased even for excisions less than 10 mm in depth, this was almost twofold for excisions of more than 10 mm in depth.
but no treatment only included three studies and 3764 women with high grade squamous intraepithelial lesion consequence of treatment but rather a product of other than 10 mm, threefold for more than 15-17 mm, and 20 mm) versus untreated women

Preterm birth (37 weeks) in women treated for CIN have a higher baseline risk of prematurity than the participants; we were unable to draw any firm conclusions from this comparison. When we assessed the risk of preterm birth according to both the cone depth and comparator, we noted overall the same direction of effect. Although the difference in the risk of preterm birth for small excisions (≤10-12 mm) compared with just having CIN but no treatment, became insignificant, the number of studies assessing that comparison was small, and we cannot draw firm conclusions.

Our results also confirm that although women with CIN have a higher baseline risk of prematurity than the general population, cervical treatment, and particularly deep cones, further increase that risk.

Strengths and limitations
This is the first systematic review to show that any local cervical treatment technique (excisional or destructive)
is associated with an increased risk of preterm birth and adverse obstetric sequelae and to document that the risk directly correlates to the cone depth (and volume), the treatment technique (excision more than ablation), and radicality. This meta-analysis included a large number of studies (71 cohorts) with sufficient sample size and power to explore several comparisons of treatment techniques and cone depths. Furthermore, we were able to perform subgroup analyses according to the comparator used and quantify the risk in different clinical groups.

The results, however, should be interpreted with caution. Because of the premalignant nature of the disease, no randomised studies could be identified. All the included studies were cohorts, nearly all retrospective. Such reports are at known risk of recall bias and inadequate adjustment for known and unknown confounders, while some of the outcomes of interest were difficult to measure objectively. Many of the studies relied on data collected from structured interviews and mailed questionnaires, and in some of these the response rate was small, also increasing the risk of incomplete outcome data (attrition) and misclassification bias. The studies often had different designs and used comparisons between and among women and mixed matching. Although the overall number of studies was large, for some outcomes and comparisons there were few studies, and the analyses did not have sufficient sample sizes to support definite conclusions.

Although heterogeneity between studies was not significant for most of the analyses, some subgroup analyses did show variation in the outcomes across studies. This was often in analyses that included small number of studies and participants. Meta-regression was possible for some but not all possible confounders. For many moderators, data were reported only in a proportion of the included studies. When these studies were not deemed representative of the whole population of studies, we did not perform meta-regression as this would introduce bias. Sensitivity and subgroups analyses based on study quality did not change the effect of the meta-analysis.

There were further limitations in the interpretation of the data. The gestational age cut-offs used for the definitions of severe and extreme prematurity and for different cone depths varied slightly across studies; we did show variation in the outcomes across studies. This was often in analyses that included small number of studies and participants. Meta-regression was possible for some but not all possible confounders. For many moderators, data were reported only in a proportion of the included studies. When these studies were not deemed representative of the whole population of studies, we did not perform meta-regression as this would introduce bias. Sensitivity and subgroups analyses based on study quality did not change the effect of the meta-analysis.

There were further limitations in the interpretation of the data. The gestational age cut-offs used for the definitions of severe and extreme prematurity and for different cone depths varied slightly across studies; we merged these in broader groups for the analysis. Individual patient meta-analysis data are required to more accurately describe the stratified risk of preterm birth for individual cone depths. The data on cone dimensions relied on retrospective data recorded in histopathology reports of formalin fixed samples, with obvious limitations. The formulas used for the calculation of volume also varied across studies. Future research...
should aim to correlate outcomes with precise prospective cone depth and vaginal measurements. Both the included and excluded studies showed a wide range of inclusion/exclusion criteria and outcome measures limiting statistical pooling of all the primary studies. There should be agreement among colposcopists and obstetricians on core research clinical outcome measures in line with the CROWN initiative of the premier reproductive health journals. This would improve the applicability of findings of primary and secondary research internationally.

Interpretation in light of other evidence

With an increasing evidence base suggesting that this risk is higher for more radical techniques, there has been a tendency to use less aggressive treatments. Although it was previously thought that the various techniques had comparable efficacy, evidence from a population study...
based study raised concerns that less radical treatment could increase the risk of invasion after treatment.\textsuperscript{13,14} Although the decreased number of hysterectomies could explain this increase, the move to less radical local conservative treatments is another plausible explanation. Additionally, since the first documentation of the reproductive risk associated with treatment almost a decade ago,\textsuperscript{1} subsequent observational studies and even meta-analyses reached contradictory conclusions\textsuperscript{2-4} and initiated debates within the scientific community. With some authors raising concerns that the progressive reduction in the radicality of treatment has led to increased risk of future of invasion,\textsuperscript{15-17} and others advocating the move to less radical techniques like laser ablation for the prevention of future perinatal morbidity and mortality associated with treatment,\textsuperscript{18} high quality synthesis of the evidence had become an urgent unmet need. Some of the previous small meta-analyses had methodological flaws and attempted analysis of individual treatment techniques or subgroups, thereby minimising the validity of their findings in context with the rest of the literature.\textsuperscript{16-18} All the published meta-analyses failed to analyse the data according to major confounders and stratifiers of risk, the comparison group, and the depth of the excision. Although Bruinsma and Quinn first approached the comparison group as a possible confounder, data on the depth and dimensions of the treatment were not available.\textsuperscript{5}

Preterm birth is a major cause of neonatal death and disability and represents an enormous cost to health services and society. While pregnant, these women with a history of cervical treatment make up a large proportion of referrals to specialist preterm labour prevention clinics. These referrals have increased from almost none in 1999, to more than 40% in 2012.\textsuperscript{19} Ultrasound directed surveillance is labour intensive, costly, and can be associated with maternal anxiety, more so because 85% of women after excision are effectively low risk and will deliver at term.\textsuperscript{20}

With rapidly accumulating evidence correlating cervical treatment to adverse reproductive morbidity, quantification of the comparative obstetric morbidity for different treatment techniques and cone depths is required to assist clinicians’ decision making and counselling. The results of our meta-analysis will allow clinicians, patients, and policy makers to balance the absolute increase in reproductive morbidity with increasing treatment radicality. Patients should be informed that treatment increases the risk of preterm birth compared with having CIN only, but the absolute increase in risk in small type 1 excisions is likely to be low, if any.

Furthermore, the quantified individual risk stratified by treatment and cone depth could allow obstetricians to select women considered to be at high risk of preterm birth who would benefit from intensive surveillance antenatally and minimise the unnecessary interventions for those at low risk. The antenatal management of women after treatment has been inconsistent and largely unit or clinician dependent.\textsuperscript{20} The risks and benefits associated with various interventions in pregnant women with a history of cervical treatment have not been fully assessed in properly designed studies.\textsuperscript{140} Future research should assess their value in this distinct clinical group and devise a logical strategy.

**Conclusion**

Women with CIN have a higher baseline risk of preterm birth than women from the general population. Local cervical treatment for preinvasive or early invasive disease further increases the risk, more so for excisional but also for ablative techniques. The risk of preterm birth increases with increasing cone depth (and volume) and techniques that remove or destroy larger parts of the cervix. The increase in risk for small excisions compared with having CIN is likely to be small, if any; more data are required.

In the decision to treat women of reproductive age, every effort should be made to perform a local treatment that will optimise the chances of a healthy pregnancy without compromising the completeness of the local treatment. Quality assurance in treatment of disease should include audit of dimensions of excisional specimens and persistent disease rates to ensure that treatment depth is kept to acceptable parameters (that is, at least 8 mm to include the crypts) and that oncological outcomes are not compromised.

Future research should investigate whether women who have preinvasive cervical disease are susceptible to both the disease and preterm birth, or whether HPV induced disease alone is the principal factor in increasing premature delivery. It is likely that a combination of immunological and other factors play a role. The uptake of prophylactic vaccination has been mixed in the developed world and minimal in low income countries. The impact of cervical treatment will continue to be relevant for many decades, and therefore robust clinical research in this field should remain a priority.

**Contributors.** The study was conceived and designed by MK, MA, PB, and EP. MK, AA, and MP acquired the data, which were collated and analysed by MK, AA, IK, and MA. All authors drafted and critically revised the manuscript for important intellectual content. MA, PB, and EP are joint senior authors. All authors gave final approval of the version to be published and have contributed to the manuscript. MK is guarantor.

**Funding:** This work was supported by the British Society of Colposcopy Cervical Pathology Jordan/Singer Award (P47773) (MK); the Imperial College Healthcare Charity (P47970) (AM and MK); Genesis Research Trust (P555459) (MK); Signid Jusélius Fellowship (P52483 to IK and MK), and the Imperial Healthcare NHS Trust NiHR Biomedical Research Centre (P45272) (MK, PRB); MA received support from the COHEAR-R Network (grant No 603019), funded by Seventh Framework Programme of DG Research of the European Commission (Brussels); Institut National du Cancer (Paris) through the COSPCC study (Conséquences obstétricales du (sur)traitement des précurseurs du cancer du col); the European Federation of Colposcopy (Birmingham), and the Joint Action CANCON. None of the funders has any influence on the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

**Competing interests:** All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coiDisclosureForm.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work. Ethical approval: Not required. Data sharing: No additional data available.

**Transparency:** The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.
This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 3.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/3.0/.


AOG.0b013e3181b5cba3.


© BMJ Publishing Group Ltd 2016