“IT’S ALL TOO MUCH” *

THE SHADOW OF OVERTREATMENT LOOMS OVER HYSTEROSCOPIC METROPLASTY FOR SEPTATE UTERUS

Paolo Vercellini, M.D. a,b ORCID 0000-0003-4195-0996
Francesca Chiaffarino, Biol.Sci.D. a ORCID 0000-0001-6009-8108
Fabio Parazzini, M.D. b ORCID 0000-0001-5624-4854

Unit, Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico, Milan, Italy
bUniversità degli Studi, Milan, Italy


Correspondence: Paolo Vercellini, M.D., Department of Clinical Sciences and Community Health,
Università degli Studi di Milano and Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico,
Via Commenda, 12 - 20122 Milan, Italy
Tel: +39.02.5503.2318; e-mail: paolo.vercellini@unimi.it

RUNNING TITLE: Hysteroscopic metroplasty for septate uterus

KEYWORDS: septate uterus, hysteroscopy, metroplasty, surgery, miscarriage, infertility

INTRODUCTION
The weather is getting stormy over hysteroscopic metroplasty for septate uterus, as the long-awaited results of the first randomised controlled trial (RCT) on septoplasty have finally been published in this issue of *Human Reproduction* (Rikken et al., 2021). Live birth within 12 months of randomisation, the main outcome measure, occurred in 12 of 39 (31%) women allocated to septum resection and in 14 of 40 (35%) women allocated to expectant management. The authors conclude “in light of the lack of any evidence of effectiveness and the potential for harm, we recommend against septum resection as a routine procedure in clinical practice” (Rikken et al., 2021).

The contrasts between gynaecologists fostering or opposing the hysteroscopic procedure already sharpened when no benefit of septoplasty on post-operative live birth rate was observed in two recent cohort studies (Rikken et al., 2020a; Whelan et al., 2020a). Indeed, whenever the results of good-quality studies contradicting existing convictions and habits are published, clinicians generally scrutinise protocol details, study conduct and data analysis in search for possible methodological flaws (Alonso Pacheco et al., 2020; Garzon et al., 2020; Rikken et al., 2020b; Saridogan et al., 2020; Whelan et al., 2020b). However, this might dismiss the strong overall message that an RCT, particularly in the surgical field, offers.

**ANATOMICAL PLAUSIBILITY AND BIOLOGICAL GRADIENT**

The arcuate uterus is not associated with adverse pregnancy outcomes. Contrarily, a partial or complete septate uterus has been associated with an increased risk of miscarriage, suggesting a direct relationship between the degree of uterine cavity bipartition and the probability of pregnancy loss (Akhtar et al. 2020; Chan et al., 2011; Practice Committee of the American Society for Reproductive Medicine, 2016; Venetis et al., 2014). When Fedele et al. (1989) studied the evolution of pregnancies exclusively in women with an uncorrected complete septate uterus, the risk of miscarriage was maximum when implantation occurred on the septum, and almost non-existent when it occurred on the lateral uterine
wall. This increase in the risk of miscarriage has been traditionally attributed to the macro- and microscopic differences observed between the normal myometrium and the non-resorbed tissue that divides the uterine cavity (Bettocchi et al., 2007; Fascilla et al., 2020). Anomalies in the endometrium covering the septum may also play a pathogenic role (Rikken et al., 2019).

More generally, if the septate uterus has an impact on pregnancy, the risk of miscarriage should be proportional to the likelihood of septal implantation. In fact, it seems unlikely that the prognosis is an all-or-none phenomenon based on precise cut-off millimetric measurements, percentages of uterine wall thickness, or degrees of the angle of septal indentation. However, the available evidence does not support such a biological gradient concept (Practice Committee of the American Society for Reproductive Medicine, 2016) and we are only left with the different classification of arcuate and septate uterus.

WHAT WAS THE ORIGINAL STUDY QUESTION AND WHAT ANSWER CAN THIS RCT PROVIDE?

As acknowledged by the authors themselves (Rikken et al., 2018 and 2021), participant recruitment was a major problem encountered during their TRUST trial. This is reflected by the long enrolment period from 2010 to 2018, change from a single- to a multi-centre study in 2015, change in the eligibility criteria from exclusively recurrent pregnancy loss to include subfertility in 2011 and one pregnancy loss only or preterm birth in 2015, and change in the classification adopted for the diagnosis of a septate uterus (Rikken et al., 2018).

Whelan et al. (2020a) maintain that the TRUST study “was not designed to definitively answer whether hysteroscopy metroplasty is of benefit in the subset of women with recurrent early pregnancy loss”, as women with a history of subfertility or preterm delivery were also included. According to the Practice Committee of the American Society for Reproductive Medicine (2016), “there is insufficient
evidence to conclude that a uterine septum is associated with infertility”, and the National Institute for Health and Care Excellence (NICE), as well the Royal College of Obstetricians and Gynaecologists (RCOG), suggest hysteroscopic metroplasty in women with recurrent miscarriage but not in subfertile women. The association with preterm birth is also controversial (Akhtar et al., 2020; Chan et al., 2011; NICE 2015a and 2015b; Practice Committee of the American Society for Reproductive Medicine, 2016; Venetis et al., 2014).

In addition, the ESHRE classification (Grimbizis et al., 2013) has been criticized for the risk of overdiagnosis and the limited discrimination potential between arcuate and subseptate uterus compared with the American Fertility Society scheme (The American Fertility Society, 1988). Accordingly, the RCOG suggests using the latter classification (Akhtar et al. 2020). If some women with borderline anomalies were also included in this RCT, the effect of hysteroscopic septoplasty might have been diluted (Practice Committee of the American Society for Reproductive Medicine, 2016).

Originally, Rikken et al.’s RCT was devised as a pilot study “to assess feasibility for a larger adequately powered trial” (https://www.isrctn.com/ISRCTN28960271; accessed on 20 February 2021). The authors calculated the sample size anticipating an improvement in live birth rate from 35% without surgery to 70% with surgery. They stated that such a difference was based on the results of available retrospective studies (Rikken et al., 2018). However, such a 35% difference was initially calculated for women with recurrent miscarriage only and, in theory, might not refer to those with a history of subfertility, a population subset that might not benefit so dramatically from hysteroscopic metroplasty (Akhtar et al., 2020; NICE 2015b; Practice Committee of the American Society for Reproductive Medicine, 2016).

Thus, the answer the TRUST trial provides is that a benefit of 35% or more in live birth rate after septoplasty can be confidently excluded in a mixed population including women with a history of
recurrent miscarriage, subfertility and preterm birth. The small numbers prevent separate analyses to
assess the specific effect of septoplasty in the above three study subgroups.

HOW LARGE SHOULD BE THE EFFECT SIZE?

Beyond statistical significance, the value of any medical intervention is the result of a complex
relationship between effect size, risk of harm and cost. The risk of harm of hysteroscopic metroplasty is
modest. Rikken et al. (2021) observed a single uterine perforation out of 39 procedures (2.6%). The
potential consequences of this complication are highly variable. In several large studies the incidence
of uterine perforation was very low or non-existent, although non-randomised studies are often
conservative in estimating absolute risks of harms (Papanikolaou et al., 2006). Even a procedure with a
demonstrable effect could be labelled as “low-value” if the magnitude of the statistically significant
incremental benefit achieved is not worthwhile to the health care resources it consumes (Pandya, 2018).
The cost of metroplasty performed with standard mechanical instrumentation and conducted under
ultrasonographic control, thus avoiding laparoscopy, is modest, as neither endotracheal intubation nor
overnight hospital stay are required.

With regard to the effect size, Rikken and co-workers (2021) set the between-group difference
at 35% and consider that, to detect a smaller improvement of 10% in live births, a new study would
need to recruit about 750 women and is thus likely unfeasible. However, how would patients consider
differences smaller than 35%, but larger than 10%? As an example, at the usual levels of alpha 0.05
and beta 0.20, 122 participants would be needed to detect a 25% difference. Such an effect size might
be considered both sufficiently large to justify the procedure, given the little resources consumed and
the small risk of harms, and not too large as to practically impede the organisation of another RCT.

FROM THE OPERATING THEATRE TO THE LABOUR WARD: BEHIND LIVE BIRTH RATE
Septate uterus is associated with malpresentation (Akhtar et al., 2020; Practice Committee of the American Society for Reproductive Medicine, 2016). This has been confirmed by Rikken et al. in both their cohort study (2020a) and RCT (2021), although the between-group difference in breech presentation rate was not statistically significant in either. However, as a simple exercise, and bearing in mind the limitation of pooling data from different study types without adjusting for potential confounders, when data on ongoing pregnancies is pulled from Rikken et al.’s cohort observational and randomised studies (2020a and 2021), breech presentation was significantly less frequent in women who underwent hysteroscopic metroplasty than in those who underwent expectant management (19% vs 33%; OR 0.43, 95% CI, 0.22 to 0.82. Figure 1). According to the RCOG (2017), breech presentation occurs in 3–4% of term deliveries, is associated with uterine abnormalities, and has a significant recurrence risk. In the Term Breech Trial, perinatal mortality, neonatal mortality, or serious neonatal morbidity was significantly lower for the planned caesarean section group than for the planned vaginal birth group (Hannah et al., 2000).

Presumably, nowadays most women would not accept to deliver vaginally when their baby is in breech presentation. Moreover, in case of abdominal delivery because of malpresentation, women with an uncorrected uterine septum would likely undergo repeat caesareans in subsequent pregnancies. On the other hand, a history of uneventful hysteroscopic metroplasty is usually not considered per se a contraindication to vaginal delivery, as the very few reported cases of uterine rupture during labour occurred solely when septal incision was excessive with penetration into the myometrium or uterine wall perforation (Practice Committee of the American Society for Reproductive Medicine, 2016).

However, estimating the magnitude of the effect of hysteroscopic metroplasty as a malpresentation-reducing procedure in women with septate uterus is problematic because the frequency of breech presentation in the intervention group of both studies conducted by Rikken et al. was unusually high (2020a and 2021). Thus, the overall surgical burden encumbered upon women with
septate uterus, if they eventually chose to undergo hysteroscopic metroplasty or expectant management, remains to be defined.

PROSPECTUS: DO WE NEED ANOTHER TRIAL?

Rikken and co-workers are to be greatly commended for finally adopting the only adequate scientific and ethical approach to disentangle an important and relatively frequent clinical issue. The community of hysteroscopists should now feel somewhat embarrassed when realising that septoplasty has been introduced into practice without sufficiently robust evidence of effectiveness, although recommendations for producing reliable data on surgical innovation were published more than a decade ago (Bakun et al., 2009; Ergina et al., 2009; McCulloch et al., 2009; McCullough and Jones, 2009). In addition, the authors confirmed that findings of RCTs are generally much less exciting than those of observational studies (Ioannidis et al., 2001; Selman et al. 2008).

Consequently, should hysteroscopic metroplasty be banned *tout court* without trying to clarify those doubts that remain? One of the main benefits of the current trial is that, from now on, septoplasty will be put under scrutiny all over the world. Contrary to the authors’ opinion, it may well turn out that, precisely because of the results of their study, surgeons will eventually become available for RCTs and women more willing to participate. In fact, considering both live births and delivery mode, equipoise still seems to exist, and data from another, larger, RCT might be warranted.

In this regard, the ESHRE Special Interest Group in Reproductive Surgery could organise an adequately powered, multi-national trial selectively recruiting women with class U2b anomalies (Grimbizis et al., 2013 and 2016); that is, complete septate uterus. Such a trial would prevent any potential classification bias. Either live birth or caesarean section rate could be selected as the primary outcome. Academic “adversaries” should participate in designing the RCT with the aim of limiting potential investigator bias, and data could be analysed by independent research groups (Leichsenring
and Steinert, 2017). Imaging and surgical findings should be rendered publicly available together with trial data.

In the meantime, women should be fully informed of all the uncertainties surrounding hysteroscopic metroplasty in different clinical conditions. The possibility that septoplasty may confer only a limited benefit, or no benefit at all, must be clearly revealed. The surgeon should disclose any financial conflict of interest in case the procedure is scheduled. Precisely when the quality of the evidence in favour of surgery is weak, different women may choose differently if adequately informed and empowered (Johnson et al., 2008).
DATA AVAILABILITY

No new data were generated in support of this manuscript. Pooled analysis was based on data reported in two published reports (Rikken et al. 2020a and 2021).

AUTHORS’ ROLES

P.V. conceived and drafted the original version of the manuscript; F.C and F.P. participated in the conception of the manuscript; all the authors revised critically the article for important intellectual content, and approved the final version of the manuscript to be published.

FUNDING

None

CONFLICT OF INTEREST

P.V., F.C., and F.P. declare that they have no conflicts of interest.
REFERENCES


Leichsenring F, Steinert C. Is cognitive behavioral therapy the gold standard for psychotherapy? The need for plurality in treatment and research. *JAMA* 2017;318:1323-1324.


Saridogan E, Mavrelos D, Jurkovic D. To decide on the value of hysteroscopic septum resection we need prospective data. Hum Reprod 2020;35:2627.


Figure 1. Forest plot showing individual and combined effect size estimates and 95% CIs in two studies that evaluated the likelihood of breech presentation in women with a subseptate or septate uterus who underwent hysteroscopic metroplasty or expectant management before the index conception. Horizontal lines indicate 95 CIs; boxes show the study-specific weight; rhombus represents combined effect size.
<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Septum resection</th>
<th>Expectant management</th>
<th>Odds Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>Total</td>
<td>Events</td>
</tr>
<tr>
<td>Rikken et al., 2020</td>
<td>17</td>
<td>89</td>
<td>27</td>
</tr>
<tr>
<td>Rikken et al., 2021</td>
<td>2</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>19</td>
<td>102</td>
<td>94</td>
</tr>
</tbody>
</table>

Total events: 33

Heterogeneity: Tau^2 = 0.00, Chi^2 = 0.42, df = 1 (p = 0.52), R^2 = 0%

Test for overall effect: Z = 2.54 (p = 0.01)