4 Reproducibility of the interpretation of pelvic X-ray 3 months after hysteroscopic sterilization with Essure.

S. Veersema
B.W.J. Mol
H.A.M. Bröllmann

Abstract

Objective To estimate the diagnostic accuracy and the interobserver reproducibility of pelvic X-rays in the diagnosis of successful bilateral sterilization with Essure after a 3-month follow-up period.

Design Interobserver study.

Setting Outpatient department of obstetrics and gynecology in a Dutch teaching hospital.

Patient(s) Patients with successful bilateral Essure placement.

Intervention(s) Hysteroscopic sterilization with Essure and pelvic X-ray and hysterosalpingography after a 3-month follow-up period.

Main Outcome Measure(s): Six observers evaluations of 47 pelvic X-rays from 47 patients 3 months after a technical successful bilateral placement of microinserts to estimate the reliability of the sterilization. Diagnostic accuracy of pelvic X-ray per observer in detecting incorrectly positioned microinserts was expressed in terms of sensitivity and specificity, with hysterosalpingography as the reference strategy. Reproducibility of the interpretation of the pelvic X-ray was expressed as κ-values.

Result(s) The sensitivity and specificity for X-rays read by gynecologists was 0.67 (95% confidence interval [CI], 0.29-0.96) and 0.79 (95% CI: 0.58-1.00) and for radiologists 1.0 and 0.5 (95% CI: 0.36-0.64). The interobserver agreement in reliability of pelvic X-ray of hysteroscopic sterilization assessment with Essure ranged from slight (k-value 0.09) for gynecologists to moderate (κ-value = 0.52) for radiologists.

Conclusion(s) Test characteristics of pelvic X-ray as the imaging technique to assess the position of the Essure microinserts and tubal patency were poor, as was the reproducibility, particularly if gynecologists performed the evaluation. We do not recommend the use of pelvic X-ray for the assessment of the positioning of microinserts after hysteroscopic sterilization. (Fertil Steril 2010;94:1202–7)

Key Words Essure, hysteroscopic sterilization, confirmation test, X-ray, interobserver reproducibility
Introduction

The Essure Permanent Birth Control System (Conceptus Inc, Mountain View, CA) is a new method of proximal tubal occlusion by hysteroscopic placement of a microinsert in the uterotubal junction (1–3). The procedure is gaining popularity because it can be performed under local or no anesthesia in the office. During hysteroscopy, the introduction device is inserted in the fallopian tube, after which the device can expand and the Essure microinsert remains in position. The Essure microinsert consists of a stainless steel inner coil, a nickel titanium alloy outer coil, and polyethylene terephthalate (PET) fibers covering the inner coil. The PET fibers induce a tissue response, which causes fibrous tissue in growth with tubal occlusion (4).

The position of the devices has to be confirmed 3 months after the procedure before the patient can rely on this permanent contraception and cease her alternative contraception. Different imaging techniques are used to document localization of the microinserts and tubal occlusion 3 months after placement. The traditional hysterosalpingography (HSG) is the only imaging method currently approved by the U.S. Food and Drug Administration (FDA). The current recommendation in countries outside the United States is to check the position and the alignment of the devices with pelvic X-ray (5). The limitations of pelvic X-ray are the risk of ionizing radiation and lack of soft tissue and tubal patency information.

Adverse events after the Essure sterilization, such as subsequent pregnancy, expulsion of the microinsert, or perforation, have been described in previous studies and are associated with incorrect placement procedures (1,3).

Earlier phase II and pivotal multicenter trials were started in 1998 and strongly advised patients to use alternative contraception for 3 months after the procedure until tubal occlusion was confirmed by HSG (1,3). Satisfactory bilateral insertion was achieved in 664 of 734 patients (90%). Satisfactory placement implied good visibility of the tubal ostia in hysteroscopy and microinserts location across the uterotubal junction, with three to eight visible expanded coils trailing into the uterine cavity. The 100% bilateral tubal occlusion after an initial satisfactory bilateral placement was confirmed by HSG. To exclude unexpected failures, a less invasive diagnostic test may be sufficient.

Two earlier reports on the results of clinical trials in Spain with hysteroscopic sterilization with Essure and the use of pelvic X-ray 3 months after a satisfactory bilateral placement (6,7) are available.

A recent analysis of 1630 women who underwent an office hysteroscopic tubal sterilization with Essure between January 2003 and June 2006 showed a successful insertion rate of 99% (8). Women were advised to use an alternative contraceptive method until a simple X-ray examination was performed at least 3 months after
the insertion. Hysterosalpingography or ultrasonography was performed when the placement was not satisfactory (more than eight or fewer than three coils remaining visible by hysteroscopy, insertion in only one tube, unclear radiologic results). None of these patients have become pregnant after confirmation at the 3-month follow-up evaluation. These results indicate that pelvic X-ray may be useful to confirm successful bilateral placement in an uncomplicated procedure. However, no data are available about the diagnostic accuracy and the reproducibility of the pelvic X-ray as a confirmation test for Essure sterilization. In cases of poor test characteristics or reproducibility, routine use of pelvic X-ray may be misguided, adding no useful additional information and wasting health-care resources.

Our study was designed to test the accuracy of pelvic X-ray compared with the standard HSG as reference test among radiologists and gynecologists for the diagnosis of successful bilateral sterilization with Essure after a 3-month follow-up period and to estimate the interobserver agreement.

**Materials and methods**

The study was conducted in the Department of Obstetrics and Gynecology of the St. Antonius Hospital, a teaching hospital in Nieuwegein, the Netherlands. The approval of the institutional review board was not considered necessary as patients were not in any way involved in the study and patient data were anonymous. Moreover, the study is part of large cohort study of 100 consecutive patients (9) that was approved by the institutional review board. All women gave their written informed consent to being part of the cohort, and to knowing that the Essure technique of sterilization is irreversible and that data of a long-term follow-up study were not yet available. Patients were eligible for this study if an adequate plain pelvic X-ray as well as a HSG were digitally available. As the quality of the X-rays was suboptimal in many cases, only 47 cases were included as having digital X-rays of optimal quality.

Hysteroscopic sterilization using the Essure system inserted by use of the standard technique was performed in an outpatient setting. Three months after the procedure, a pelvic X-ray was performed with the patient in supine position, followed by a HSG with a water-soluble contrast medium (Telebrix-Polyvidone; Guerbet SA, Villepinte, France). The contrast medium was instilled into the uterine cavity after the pelvic X-ray was made, using a silicone balloon HSG catheter (Cook Ireland Ltd., Limerick, Ireland). All images were digitally recorded to enable digital demonstration afterward (5). The captured images were evaluated by a radiologist and gynecologist using the algorithm from the HSG protocol in the manufacturer’s physician training manual (5). If there was a satisfactory position of both devices
and bilateral tubal occlusion, the sterilization was considered successful, and the patient was advised to cease alternative contraception.

In 2005, all X-rays were evaluated simultaneously by an international panel formed by six observers. The observers were not informed about the clinical data of the hysteroscopic sterilization procedures. The only clinical information available at the time of evaluation was that a bilateral hysteroscopic sterilization had been successfully performed 3 months before the X-ray. The participants were three gynecologist, all specialists in hysteroscopic sterilization with good experience in reading X-rays after Essure sterilization (Gyn 1-3), two radiologists, with good and moderate experience (Rad 1 and 2), and a registrar in radiology with experience in reading HSGs after Essure sterilization (Rad R).

The observers were blinded for the results of the HSG. They were not allowed to discuss the results. The Essure X-ray protocol from the manufacturer’s physician training manual (5) was used while evaluating of the pelvic X-rays.

According to the manufacturer’s protocol, the observers evaluated the pelvic X-rays with regard to the position of the microinserts, the symmetrical appearance of the devices, and the distance between the two devices. In addition, they had to judge the X-ray as satisfactory, uncertain, or unsatisfactory and determine whether the patient could rely on the sterilization (Table 1).

Table 1
Observer assessment items for 3-month evaluations of Essure placement X-rays.

<table>
<thead>
<tr>
<th>Evaluation items</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to assess</td>
<td>Good</td>
</tr>
<tr>
<td>Position right device</td>
<td>Correct</td>
</tr>
<tr>
<td>Position left device</td>
<td>Correct</td>
</tr>
<tr>
<td>Symmetrical appearance</td>
<td>Yes</td>
</tr>
<tr>
<td>Distance</td>
<td>&lt;4 cm</td>
</tr>
<tr>
<td>Conclusion</td>
<td>Satisfactory</td>
</tr>
<tr>
<td>Rely on Essure?</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Notes: The observers evaluated the pelvic X-rays on ability to assess the position of the microinserts, symmetrical appearance of the devices, and the distance between the two devices. They also had to grade the X-ray as satisfactory, uncertain, or unsatisfactory and determine whether the patient could rely on the sterilization.
Statistical Analysis

Reproducibility was expressed using Fleiss’s $\kappa$-statistics (10). Fleiss’s kappa ($\kappa$) works for any number of raters giving categorical ratings to a fixed number of items. It can be interpreted as expressing the extent to which the observed amount of agreement among raters exceeds what would be expected if all raters made their ratings completely randomly. A $\kappa$-value of 0 indicates no agreement beyond chance, a $\kappa$-value of 1 indicates perfect agreement between observers. The reproducibility in the case of $\kappa$-values between 0 and was regarded as slight, between 0.2 and 0.4 as fair, between 0.4 and 0.6 as moderate, between 0.6 and 0.8 as substantial, and between 0.8 and 1.0 as almost perfect.

The $\kappa$-values were calculated for the six observers together and for the three gynecologists and three radiologists separately.

In the assessment of diagnostic accuracy of the X-ray, HSG was considered to be the reference test.

Diagnostic accuracy was calculated for each observer and expressed in terms of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). An unsatisfactory or uncertain evaluation of an X-ray as of the position of the microinserts was regarded as a positive test result, whereas a satisfactory evaluation was regarded as a negative test result (successful sterilization).

Results

Between December 2003 and July 2004, 47 patients with technically successful bilateral placement of Essure were included in the study. All patients were evaluated by HSG and pelvic X-ray after a 3-month follow-up period. In 44 cases, the HSG confirmed correct position of both implants and bilateral occlusion of the fallopian tubes. There were three cases of an abnormal position of one of the microinserts and a patent tube on HSG: one complete expulsion, one expulsion to the uterine cavity (Fig. 1), and one perforation with the device in the abdominal cavity (Fig. 2).

The diagnostic accuracy for each observer is expressed in terms of sensitivity and specificity. The pooled sensitivity and specificity of both gynecologists and radiologists was 15 out of 18 (83%) and the specificity 170 out of 264 (64%). According to medical specialty, the sensitivity for the three radiologists was 100%, and the specificity was 66 out of 132 (50%). For 27 cases, at least one of the radiologists advised additional HSG to confirm a reliable sterilization (false positive). The sensitivity for the X-ray evaluated by the gynecologists was 6 out of 9 (67%), and the specificity was 79%. Two gynecologists accepted the X-ray from the case with the perforation as satisfactory, and one gynecologist did not recognize the expulsion into the cavity (false negative) (Table 2).
The overall interobserver agreement with regard to reliability of the hysteroscopic sterilization with Essure ranged from slight ($\kappa=0.09$) for gynecologists to moderate ($\kappa=0.52$) for radiologists (Table 3).

There was agreement between all observers that the sterilization was satisfactory in 11 cases, with advice to the patients that they could rely on the sterilization. Only for the case with one patent microinsert (due to expulsion of the other) was there agreement between all six observers that the sterilization was unsatisfactory. The agreement between the observers on the visibility of both devices was perfect, while the agreement on the position of the devices ($0.28$ to $0.30$), the symmetrical appearance of both devices ($0.37$), and the distance between the two devices ($0.27$) was fair. The agreement on the final conclusion of the X-ray was slight ($0.17$) (Table 3).

The patient with the perforation underwent a laparoscopic sterilization with Filshie Clips (Femcare-Nikomed Limited, Hampshire, UK.). During the same procedure, the microinsert was released from the omentum. The two other patients with incorrect position of the microinsert underwent a second successful hysteroscopic sterilization confirmed by HSG after 3 months. None of the patients has become pregnant.
Table 3
Interobserver agreement kappa values.

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Radiologist</th>
<th>Gynecologists</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>κ  95% CI</td>
<td>κ  95% CI</td>
<td>κ  95% CI</td>
</tr>
<tr>
<td>Right device</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visible</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Optimal position</td>
<td>0.28</td>
<td>0.25–0.31</td>
<td>0.45*</td>
</tr>
<tr>
<td>Left device</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visible</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Optimal position</td>
<td>0.30*</td>
<td>0.27–0.33</td>
<td>0.54*</td>
</tr>
<tr>
<td>Accessibility</td>
<td>0.14*</td>
<td>0.11–0.18</td>
<td>1.0</td>
</tr>
<tr>
<td>Symmetrical</td>
<td>0.37*</td>
<td>0.34–0.41</td>
<td>0.77*</td>
</tr>
<tr>
<td>Distance</td>
<td>0.27*</td>
<td>0.24–0.30</td>
<td>0.44*</td>
</tr>
<tr>
<td>Conclusion</td>
<td>0.17*</td>
<td>0.14–0.20</td>
<td>0.25*</td>
</tr>
<tr>
<td>Rely on</td>
<td>0.24*</td>
<td>0.21–0.28</td>
<td>0.52*</td>
</tr>
</tbody>
</table>

Notes: Agreement between the three radiologists and three gynecologists about visibility, optimal position of the device, accessibility of the X-ray, symmetrical appearance, and distance between the two devices, the final conclusion for the X-ray, and the reliability of the sterilization. The reproducibility in the case of κ-values between 0 and 0.2 was regarded as slight, between 0.2 and 0.4 as fair, between 0.4 and 0.6 as moderate, between 0.6 and 0.8 as substantial, and between 0.8 and 1.0 as almost perfect. (*P< .0001.)

Discussion

In this study, six observers evaluated 47 X-rays, including three cases of incorrect position of a device. The test characteristics of the X-rays were better in the hands of radiologists (sensitivity 100%, specificity 50%) than in the hands of gynecologists (sensitivity 67%, specificity 79%). The interobserver agreement (κ) in visualizing the microinsert was 100% in both radiologists and gynecologists; however, in scoring reliability of the Essure sterilization there was a large difference in the agreement between radiologists (52%) and gynecologists (9%).

In the United States, hysterosalpingography (HSG) is the only imaging method currently approved by the FDA for the diagnosis of tubal occlusion after the Essure sterilization procedure. In other countries, other diagnostic tools are used for confirmation. In Europe, pelvic X-ray is recommended by CE Mark guidelines.
To the best of our knowledge, ours is the first report on the diagnostic accuracy of pelvic X-ray in the assessment of the correct placement of the Essure microinserts using HSG as a reference test. Hysterosalpingography is the best available diagnostic test to assess the efficacy of the Essure sterilization in terms of position of the microinserts and blockage of the tubes. However, the combination of position and patency is crucial, as blocked tubes alone do not guarantee an effective sterilization.

As to the different performance of gynecologists and radiologists in favor of the latter, any explanation is speculative. The pelvis X-rays as evaluated by the gynecologists gave a low sensitivity and evaluated by the radiologists gave insufficient specificity.

**Figure 1**
Pelvic X-ray and hysterosalpingography of patient with partial expulsion of right device into the uterine cavity.

**Figure 2**
Pelvic X-ray and hysterosalpingography of patient with perforation of the left device to the abdominal cavity.
In other words, not one radiologist confirmed an adequate sterilization for the three patients with a failure, but they advised additional HSGs to perform, whereas the gynecologists accepted more X-rays as reliable sterilizations, but two gynecologists did not recognize the perforation, and one of them did not recognize the expulsion to the uterine cavity.

Radiologists are used to evaluating images with little or no clinical data, and they are aware of the clinical consequences if they miss an important abnormality. As imaging specialists, they may be more inclined to adhere to given instructions on diagnostic criteria than doctors in other specialties. All the observers were instructed on the criteria of failed position of the microinserts according to the protocol (see Table 1).

Unfortunately, the alignment of the fourth marker, which is considered to be an important criterion nowadays, was not included in the criterion list. As the frequency of failed insertion of the Essure device is low, so is the experience of the observers.

In our case series, in three out of 47 X-rays the displacement of the microinsert was reported. This low frequency is in line with former publications (1,3,8,11). In the phase II study (1) of 226 patients evaluated with HSG after a three-month follow-up period, six perforations of the uterine wall or tubal lumen and one expulsion were reported. In the published European series with X-ray as the confirmation test, there was a 100% success rate after bilateral placement, with no unsatisfactory device locations (12, 13). In the most recent study by Arjona et al. (8), three pregnancies occurred during the first 3 months after 1650 procedures.

The issue of better instruction and training of the X-ray reviewers was also addressed by van der Leij et al. (12), who reported in 1997 on the interobserver and intraobserver agreement in the evaluation of radiographic images after hysteroscopic sterilization with Ovabloc (formed-in-place intratubal silicone devices; European Medical Contract Manufacturing B.V., CH Nijmegen, the Netherlands). A group of eight gynecologists had only poor interobserver agreement on the reliability of the sterilization. The investigators concluded that this underlined the need for training in standardized interpretation of X-rays concerning the reliability of sterilization (12).

Another reason for the poor diagnostic performance and agreement among gynecologists in particular may be the lack of clinical data, such as information on any difficulties during the Essure insertion procedure. In daily practice, these data may alert the X-ray observer that there may be cause for an underestimation of the accuracy of the X-rays.

In a recent study in the United Kingdom to determine patient satisfaction of outpatient female sterilization, the majority of women (96%; 95% CI: 88-99%) reported satisfaction with their overall experience of the Essure hysteroscopic sterilization procedure and follow-up evaluations. Two patients declined to have their scheduled HSG, and only 72% of the patients reported the HSG as an “acceptable test.”
The use of pelvic X-ray, which is less invasive and causes less inconvenience, demonstrates intra-abdominal localization of the microinserts. Because of the lack of soft tissue detail and no filling of the uterine cavity with contrast dye, no information is available regarding the relationship of the microinsert to the uterine cornua.

Our study shows that a correct position of the microinserts in the fallopian tube is difficult to assess by plain X-ray alone. Pelvic X-ray was a perfect diagnostic tool to confirm a complete expulsion of a microinsert, but it missed discrete dislocation where the dislocated microinsert was still attached to the uterine wall. The lack of additional clinical information and history of the patient, and possibly the insufficient training of the observers may be associated with the poor test characteristics and reproducibility of pelvic X-rays after Essure sterilization. The results of our study do not justify the routine use of this radiographic tool in clinical practice.

Recently, it was shown (13-17) that transvaginal ultrasound assessment may be as reliable as HSG for uterotubal localization of the microinserts. The use of ultrasound obviates the need for ionizing radiation in the majority of patients. Future studies on other imaging techniques such as ultrasound are necessary to optimize imaging after Essure placement.

Acknowledgments: The authors thank L. Koobs, A. Ubeda-Hernandez, Professor H. Fernandez, Professor J. Cordon, R. Peck, and W. van Es for the interpretation of the pelvic X-rays and J. C. Kelder for the help with the analysis of the data.
References