Uterine artery embolization: Long term follow-up and implementation
van der Kooij, S.M.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
The effect of treatment preference and allocation on patients’ health related quality of life in the randomized EMMY trial

SM van der Kooij
WJK Hehenkamp
E Birnie
WM Ankum
BW Mol
S Scherjon
JA Reekers

ABSTRACT

Objectives
To determine the effect of preference and treatment allocation on health related quality of life (HRQOL) in patients in the randomized EMMY trial of hysterectomy versus uterine artery embolization (UAE) for symptomatic uterine fibroids.

Study design
We invited 349 patients eligible for trial participation, of which 177 agreed to participate (the ‘randomized group’). Within the randomized group patients were allocated to UAE (n=88) or hysterectomy (n=89). The remaining 172 patients refused randomization and received the treatment of their preference (varying from hysterectomy to no treatment at all), of which 103 patients agreed to fill in questionnaires (the ‘preference group’). Patients’ treatment preferences and HRQOL were assessed at baseline and the patients were prospectively followed to evaluate HRQOL 12 months after treatment.

Results
At baseline, most patients in the randomized group preferred UAE: 115/177 (65%). In the preference group most patients preferred hysterectomy: 100/172 (58%). At 12 months there was no effect of having had the preferred treatment on HRQOL, neither in the randomized nor in the preference group. The randomized group improved significantly in both mental and physical health, compared to baseline. In the preference group only mental health improved compared to baseline, while physical health did not improve significantly.

Conclusions
In a randomized trial comparing UAE and hysterectomy for symptomatic fibroids, the pre-randomization preference for a specific treatment did not affect HRQOL.
Trial participants improved better on physical HRQOL than women who refused to participate.
INTRODUCTION

In the assessment of new interventions, the randomized controlled trial (RCT) is generally acknowledged as being the optimal study design when evaluating treatment effectiveness, but the design of an RCT does not allow room for patients’ treatment preferences, which may influence the outcome of Health Related Quality of Life (HRQOL)[1]. Firstly, RCT participants may have a strong preference for a ‘new’ intervention, if this new intervention is perceived as being more effective and/or less invasive. This may cause selective participation bias. Also, there may be ‘preference effects’, where outcomes may differ depending on whether or not an individual is randomized to his/her preferred treatment [1]. This effect occurs especially in trials where patients cannot be blinded for the allocated treatment and when outcome measures are subjective, e.g. HRQOL [2]. We compared pre-procedural preferences in an RCT with an obvious difference in the level of invasiveness between the two treatment arms. The data we used for the present study are part of the randomized controlled EMMY trial (EMbolization versus hysterectoMY) comparing two treatment strategies for symptomatic uterine fibroids [3]. Within that trial we demonstrated the minimally invasive treatment arm, i.e. uterine artery embolization (UAE), to be a good alternative for hysterectomy, a major surgical intervention. The aim of the present study is to explore the impact of patient preferences in a randomized controlled trial to see if actually taking part in the RCT had an impact on HRQOL that was different from deciding on the treatment to be undertaken in advance and refusing to participate. In addition we looked at the HRQOL of those women who had been randomized who might have preferred to have the treatment other than the randomly allocated one.

MATERIALS AND METHODS

The EMMY trial

The EMMY trial is a multi-centre RCT, conducted in the Netherlands. Patients visiting the gynaecological outpatient clinics were invited to participate if they were pre-menopausal; were diagnosed with uterine fibroids; had menorrhagia; had no other treatment option than a hysterectomy and had no desire for future pregnancy. When patients agreed to participate they were randomized 1:1 to either hysterectomy or UAE. The trial design and results are described in more detail elsewhere [3].
Patients

All patients that agreed to be randomized are defined in this paper as the ‘randomized group’. They were treated according to randomized treatment allocation. All eligible patients who declined trial participation for various reasons comprised the ‘preference group’. These women were treated according to their preference. Patients in the preference group who underwent another treatment than UAE or hysterectomy, formed the ‘other treatment’ subgroup. At the time of the study UAE was largely unavailable outside the trial in the Netherlands.

Treatment information

All patients eligible for trial participation were informed both verbally and in writing prior to randomization about the risks and benefits of both procedures, based on data from the literature available at study onset (2002).

Hysterectomy was performed mostly abdominally and described as a procedure performed under general anaesthesia with 5-7 days in-hospital stay; a recovery time of about 6 weeks; a 100% success rate for solving heavy menstrual bleedings; and an unspecified but low risk of complications such as severe infection, premature ovarian failure or urinary incontinence. UAE was described as a minimal invasive technique, performed under local or epidural anaesthesia with a 1-2 days in-hospital stay; recovery on average within 2 weeks; about 80% likelihood of resolving menorrhagia complaints; a uterine volume reduction of about 50%; and an unspecified risk of infection, premature ovarian failure or need for additional surgical intervention.

Procedures

UAE and hysterectomy were performed according to professional standards as described in the trial report [3]. Patients in the preference-group were not randomized, but received the treatment of their own choice.

Questionnaires

All eligible patients from both the randomized and the preference group were asked to complete a questionnaire at baseline, i.e. before randomization or treatment selection. Patients in the randomized group filled out a follow-up questionnaire at 12 months. Patients in the preference group received a survey by mail inquiring which treatment(s) they had received and when this intervention had taken place. Next, they were invited to complete...
the follow-up questionnaire 12 months after intervention. Patients in the preference group who received no treatment were also invited to complete the follow-up questionnaire at 12 months after the baseline questionnaire.

The baseline questionnaire consisted of two parts; 1) treatment preference: preference for hysterectomy, for UAE or being indifferent between hysterectomy and UAE; and 2) the Medical Outcome Study Short Form-36 (SF-36), a validated general quality of life questionnaire which can be summarized into a mental health component (MCS) and a physical health component (PCS) summary score [4]. The follow-up survey after 12 months consisted of the SF-36 questionnaire.

**Statistical analysis**

Data entry was performed using SPSS data entry for Windows 3.0. Analyses were done using SPSS statistical software (version 11.5.1). Study outcomes within the randomized group were analyzed according to original treatment assignment (intention to treat). Baseline characteristics of the randomized group were compared to those of the preference group by logistic regression analysis. In the first comparison within the randomized group the following statistics were used. Differences in pre-procedural preferences were expressed in n (%). Differences in categorical data between the groups were compared with x2-tests or Fisher Exact tests if appropriate. Differences in paired categorical data within patients were tested with Wilcoxon matched pairs test. Multiple binary logistic regression analysis was used to investigate the effect of various co-variables on baseline preferences: hysterectomy versus UAE, disregarding indifference. Co-variables entered in the multivariable model were those variables yielding p-values <0.1 in the univariate analysis (appendix 1). Multiple linear regression analysis and multiple logistic regression analysis were used to investigate whether the agreement between allocated treatment and baseline treatment preference was of any influence on the SF-36 PCS and MCS change scores (compared to baseline). Entered in the univariate analysis were: treatment match (undergoing preferred treatment/undergoing nonpreferred treatment) and baseline preferences (having any preference/having no preference). In the second comparison (randomized versus preference group) the following analyses were done. The mean scores of the SF-36 PCS and MCS at baseline and the mean change scores at 12 months follow-up were compared univariately between the randomized and the preference group by means of a Student’s t-test. Analysis of the PCS and MCS change scores was only based only on the data of patients that completed both the baseline as well as the 12 months follow-up questionnaire. Baseline characteristics (appendix 2) were included for multiple linear regression analysis with PCS and MCS as dependent variable whenever univariate analysis revealed p-values <0.1, including the
variable treatment received (hysterectomy/UAE/other surgical options/GnRH analogues/alternative medicine/no treatment). A similar analysis was performed for the identification of parameters that influenced the SF-36 change score at 12 months. A two-sided p-value of <0.05 was considered statistically significant in all other analyses.

RESULTS

Patients

Patients were enrolled between March 2002 and February 2004. Figure 1 shows the flow of patients through the study. Of 349 eligible patients, 177 were randomized: 88 were allocated to UAE and 89 to hysterectomy. These patients formed the ‘randomized group’. In the randomized group 21 patients (12%) withdrew from the trial. For follow-up, 156 randomized patients participated and 155/156 (99%) participating patients completed the baseline questionnaire.

The remaining group of 172 patients who refused participation formed the ‘preference group’. In this group 103/172 (60%) patients completed the baseline questionnaire; the 69 non-responders were excluded because of lack of information. Concerning baseline patients characteristics, these non-responders were not different from patients in the preference-group who completed the baseline questionnaire. Patients in the preference group received the treatment of their own preference. Patients in the preference group who underwent another treatment than UAE or hysterectomy formed the ‘other treatment’ subgroup.

Figure 1. Patients flow chart
The 12 months follow-up questionnaire was completed by 151/156 (97%) patients in the randomized group, and by 55/103 (53%) patients in the preference group. Baseline and procedural characteristics of the randomized and preference group are shown in table 1. The univariate analysis showed that baseline characteristics between randomized and preference group differed significantly in terms of ethnicity, i.e. more Caucasian patients in the preference group, and mental health, i.e. higher SF-36 MCS baseline scores in the preference group (Table 1).

**Table 1. Baseline characteristics.**

<table>
<thead>
<tr>
<th></th>
<th>Randomized group</th>
<th>Preference group</th>
<th>Randomized group versus Preference group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UAE N= 81</td>
<td>Hysterectomy N=75</td>
<td>UAE and Hysterectomy N=103</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>44.6 (4.8)</td>
<td>45.4 (4.2)</td>
<td>44.9 (4.9)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>24 (27.3)</td>
<td>20 (22.5)</td>
<td>13 (12.6)</td>
</tr>
<tr>
<td>White</td>
<td>54 (61.4)</td>
<td>57 (64.0)</td>
<td>82 (79.6)</td>
</tr>
<tr>
<td>Other</td>
<td>10 (11.4)</td>
<td>12 (13.5)</td>
<td>8 (6.2)</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>16 (18.2)</td>
<td>13 (14.8)</td>
<td>15 (14.6)</td>
</tr>
<tr>
<td>Married</td>
<td>55 (62.5)</td>
<td>57 (64.0)</td>
<td>72 (69.9)</td>
</tr>
<tr>
<td>Divorced</td>
<td>12 (13.6)</td>
<td>15 (17.0)</td>
<td>7 (6.8)</td>
</tr>
<tr>
<td>Widower</td>
<td>0 (0)</td>
<td>2 (2.3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>68 (77.3)</td>
<td>69 (78.4)</td>
<td>78 (75.7)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>20 (22.7)</td>
<td>19 (21.6)</td>
<td>23 (22.3)</td>
</tr>
<tr>
<td>Educational level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary school</td>
<td>3 (3.4)</td>
<td>6 (6.9)</td>
<td>1 (1.0)</td>
</tr>
<tr>
<td>Lower secondary school</td>
<td>29 (33.0)</td>
<td>32 (36.8)</td>
<td>29 (28.2)</td>
</tr>
<tr>
<td>Intermediate/high secondary school</td>
<td>26 (29.5)</td>
<td>27 (31.0)</td>
<td>26 (25.2)</td>
</tr>
<tr>
<td>College/university</td>
<td>28 (31.8)</td>
<td>22 (25.3)</td>
<td>41 (39.8)</td>
</tr>
<tr>
<td>MOS SF-36 MCS score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>41.2</td>
<td>41.3</td>
<td>46.1</td>
</tr>
<tr>
<td>MOS SF-36 PCS score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>43.7</td>
<td>43.0</td>
<td>42.9</td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hysterectomy</td>
<td>4 (4.9)</td>
<td>75 (100)</td>
<td>46 (44.7)</td>
</tr>
<tr>
<td>UAE</td>
<td>77 (95.1)</td>
<td>0 (0)</td>
<td>9 (8.7)</td>
</tr>
<tr>
<td>Other surgical options</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>7 (6.8)</td>
</tr>
<tr>
<td>GnRH analogues</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>5 (4.9)</td>
</tr>
<tr>
<td>Alternative treatment</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>2 (1.9)</td>
</tr>
<tr>
<td>No treatment</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>6 (5.8)</td>
</tr>
<tr>
<td>Withdraw/Unknown</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>28 (27.2)</td>
</tr>
</tbody>
</table>
Patients’ preferences in the randomized and preference group at baseline

Of the patients randomly allocated to hysterectomy, 52/89 (63%) preferred UAE, 10/89 (12%) were indifferent, and 20/89 (24%) preferred hysterectomy. Baseline preference was unknown in 7/89 (8%). Of the patients randomly allocated to UAE, 63/88 (75%) preferred UAE, 8/88 (10%) were indifferent and 13/88 (15%) preferred hysterectomy. Baseline preference was unknown in 4/88 (5%). There were no significant differences in distribution of these numbers between these randomized groups (p = 0.26 for hysterectomy versus UAE versus indifference; p = 0.12 for hysterectomy versus UAE).

In the preference group, 100/172 (58%) women preferred hysterectomy and 36/172 (21%) preferred UAE. The remaining group (21%) had other preferences (waiting n = 13; alternative medicine n = 4; other surgical options n = 5; GnRH analogues n = 4; unknown n = 10). Compared to the patients in the randomized group, there were significantly less patients who preferred UAE among the preference-group (115/177 [65%] vs. 36/172 [21%]; p < 0.0001).

Two baseline co-variables were predictive of baseline preference in the randomized group. Patients with better physical health (higher SF-36 PCS score) were more enthusiastic about UAE (OR: 1.08 per point; p = 0.006). Furthermore, patients with a larger uterine volume (OR: 0.998 per cm³; p = 0.03) at baseline were more likely to prefer hysterectomy.

Health related quality of life in the randomized group compared to the preference group at baseline

At baseline the preference group showed a significantly better mental health than the randomized group (SF-36 MCS 46.1 versus 41.1; p < 0.0001). Physical health did not differ significantly between both groups at baseline (SF-36 PCS preference group: 42.9 versus randomized group: 43.7; p = 0.49). In the preference group no significant differences were found between the patients that had received a hysterectomy/UAE and the ‘other treatment’ subgroup with regard to the baseline SF-36 MCS (mean difference: -1.48; 95%CI -4.95 to 1.99; p = 0.40) and PCS (mean difference: -0.54; 95%CI -4.44 to 3.36; p = 0.79). Also in the multiple linear regression analysis we showed that the randomized group had a lower mental health (β = -4.28; 95%CI -6.77 to -1.79; p = 0.001) compared to the preference group.
Health related quality of life in the randomized and preference group at 12 months follow-up

Figure 2 shows the longitudinal changes in PCS and MCS for both the randomized group and the preference group. The randomized group improved significantly in both mental and physical health, compared to baseline (SF-36 MCS: +7.85; p < 0.0001; SF-36 PCS: +4.60; p < 0.0001). In the preference group, however, only mental health improved compared to baseline, while physical health did not improve significantly (SF-36 MCS: +7.22; p < 0.0001; SF-36 PCS: +1.89; p = 0.25).

Between the randomized group and the preference group the SF-36 MCS change-scores from baseline to 12 months follow-up did not differ (mean difference SF-36 MCS change-score: -2.62; 95%CI -6.23 to 0.98; p = 0.15). The randomized group showed a larger improvement in physical health than the preference group (mean difference SF-36 PCS change score: 5.97; 95%CI 2.32 to 9.61; p = 0.0001). At 12 months follow-up no differences were found between the UAE/hysterectomy group and the ‘other treatment’ subgroup in the preference group for the SF-36 MCS change-score (mean difference: +3.6; 95%CI -3.78 to 11.04; p=0.33) and SF-36 PCS change-score (mean difference: +1.19; 95%CI -4.92 to 7.31; p=0.70). Multiple regression analysis revealed that two factors were associated with improved physical health after treatment: having participated in the randomized trial (SF-36 PCS change score: β=6.28; 95%CI 2.93 to 9.63; p=0.002) and having received hysterectomy as treatment (β=3.58; 95%CI 0.61 to 6.55; p=0.02).

Finally, in the randomized group 59/72 (82%) UAE assigned patients and 19/63 (30%) hysterectomy assigned patients received the treatment that matched their baseline preference.
There was no effect of actually undergoing the preferred treatment in comparison to the non-preferred treatment on HRQOL (mean difference in SF-36 PCS change score: -1.02; 95%CI -4.07 to 2.02; p=0.51; mean difference in SF-36 MCS change score: 0.85; 95%CI -3.67 to 5.37; p=0.71).

**COMMENTS**

We found that in the randomized group, patients with a large uterus were more likely to prefer hysterectomy (thus arguably patients with more complaints), while patients with better physical health (possibly fewer complaints) were more enthusiastic about UAE. This might indicate that patients with heavier complaints were more inclined to prefer the definite solution, i.e. hysterectomy, whereas patients with lesser complaints preferred UAE.

Nevertheless we found that despite the difference of invasiveness of the two treatment arms, (dis-)agreement between allocated and preferred treatment at baseline had no impact on the (in both groups significantly improved) HRQOL after 12 months in the randomized group, a finding corroborated by others [5,6,7].

In the preference group, physical health did not improve significantly, while the randomized group showed a significant improvement in physical health at 12 months follow-up. The so-called ‘Hawthorne effect’ might play a role here. This is a mechanism described in humans, where subjects of an experiment change their behaviour, simply because they are being studied. Extra attention given to the patient by doctors and nurses might partly explain this mechanism [8]. This effect might positively bias effects, especially in randomized groups.

A recent Cochrane review had a different conclusion: participation in RCTs is associated with outcomes similar to receiving the same treatment outside RCTs [9]. This implies that groups inside and outside the trial are comparable. When groups inside and outside a trial differ, we might not expect similar outcomes.

In our study, differences between randomized and non-randomized patients might have occurred despite the fact that patients in the preference group met the inclusion criteria for trial entrance. We found a significantly higher preference for UAE than for hysterectomy at baseline in the randomized group, whereas the majority of patients in the preference group preferred a hysterectomy (58%). This suggests a selective trial participation bias; patients with a strong preference for UAE were more likely to be randomized because UAE was hardly available outside the trial. Also, patients with heavy complaints chose for a definite solution (i.e. hysterectomy) outside the trial to make sure that they would indeed have a hysterectomy. Patients with less serious complaints probably were not ready yet for undergoing hysterectomy, as reflected by the wide variety of treatments they received, and some patients received no
treatment at all. The patients who were entered in the trial may have been those with less outspoken preferences and more ‘moderate’ complaints; an intermediate-group. They had a preference but agreed with being randomized, while (because) most of them preferred the ‘new’ intervention UAE.

Besides this, we see a slight difference in the amount of women who had college/university education between the two groups; the women in the preference group have been higher educated (although not significant). Apparently more highly educated women are more self-confident, and conscious of what they want: they want to choose the treatment themselves. Despite the non-improvement of physical health in the preference group, mental health increased significantly and even more than in the randomized group. This might be attributed to the fact that these women indeed received the treatment they preferred. This probably adds to the feeling of higher self-esteem [1], which may be regarded a biasing effect from the psychological phenomenon of ‘cognitive dissonance reduction’ [10]; a patient’s tendency to believe that the treatment selected was indeed the best option despite contrary information or experience, thus resulting in an improved mental health experience.

There are some limitations to our study. Firstly, the best design to study the effects of preferences would have been randomization between treatment of choice versus treatment allocation by chance. To our knowledge, such a study design has never been applied and was certainly not the design of our study as re-intervention and not the effect of preference was the primary outcome measure.

This may be problematic from a strictly scientific viewpoint, but obviously reflects reality in studies as ours, where the difference between a surgical procedure performed by the gynaecologist under general anaesthesia and a less invasive procedure performed by the intervention radiologist under local anaesthesia, are ingredients for women making treatment choices and preferences. Secondly, the questionnaire response-rate was significantly lower in the preference group compared to the randomized group, and there was an uneven distribution in the preference group between UAE and hysterectomy.

In conclusion, patients allocated to their preferred treatment in our randomized trial had equally good post-procedural HRQOL change scores compared to those patients who were not allocated to their preferred treatment. Patients who refused to participate in our trial, despite meeting its inclusion criteria seem to have had less to gain in terms of physical HRQOL. This may be attributed to the fact that patients inside a trial receive more attention (the Hawthorne effect), or be explained by selective trial participation bias. We think that selective trial participation bias can be a drawback for the external validity of RCT results, influencing the composition of the study-group, which, as a result, might not be that perfectly comparable to the normal population as generally believed.
REFERENCES


10. Shultz TR, Lepper MR. Cognitive resonance reduction as constraint satisfaction. Psychological Review
APPENDIX 1
Effect on baseline preference
- Age (continuous)
- Ethnicity (Caucasian/non-Caucasian)
- BMI (continuous)
- Parity (parous/non-parous)
- Smoking (yes/no)
- Co-morbidity (yes/no)
- Educational level (intermediate/higher level vs lower level)
- Previous surgical treatment (yes/no)
- Any previous treatment (yes/no)
- Duration of menorrhagia symptoms (continuous)
- Hemoglobin level (continuous)
- Anemia before treatment (yes/no)
- Number of fibroids (continuous)
- Uterine volume (continuous)
- Baseline SF-36 MCS and PCS (continuous)

APPENDIX 2
Effect on SF-36 at baseline and follow-up
- Age (continuous)
- Ethnicity (black/Caucasian/other)
- Marital status (single/married/living apart together-divorced-widow)
- Employment status (employed/unemployed)
- Education level (elementary school/lower secondary school/-intermediate-high second school/college-
  university)
- Participation status (randomized/preference)