Long-term consequences of subtotal and total abdominal hysterectomy

Follow-up of a randomized clinical trial and an observational study

Lea Laird Andersen, MD
Academic supervisors

Associate Professor, Helga Gimbel MD, DMSc

Department of Obstetrics and Gynecology Nykøbing Falster hospital and the University of Southern Denmark, Institute for Regional Health Research

Professor, Bent Ottesen MD, DMSc

Department of Gynaecology, Juliane Marie Centret, Rigshospitalet

Head of Department, Lars Alling Møller MD, PhD

Department of Obstetrics and Gynaecology, Roskilde Hospital

Assessment committee

Chairperson

Associate Professor, Pernille Ravn, MD, DMSc

University of Southern Denmark, institute for clinical research

Professor Ian Milsom MD, MB, ChB, PhD

Department of Obstetrics and Gynaecology, Göteborg University

Associate Professor, Charlotte Møller MD, PhD

Department of Obstetrics and Gynaecology, Aarhus University Hospital
Table of contents

Table of contents ........................................................................................................................................... 3
Abbreviations .................................................................................................................................................. 6
Preface .............................................................................................................................................................. 7
1. Aim ........................................................................................................................................................... 9
2. Summaries ................................................................................................................................................ 12
   2.1 English summary ................................................................................................................................. 12
   2.2 Dansk Resumé ...................................................................................................................................... 14
3. Background .............................................................................................................................................. 16
   3.1 History: ............................................................................................................................................... 16
   3.2 Hysterectomy in Denmark .................................................................................................................. 16
   3.3 Subtotal versus total hysterectomy .................................................................................................... 17
   3.4 Cervical cancer screening .................................................................................................................. 19
4. Aim .......................................................................................................................................................... 21
5. Setting ...................................................................................................................................................... 22
6. Methods ................................................................................................................................................... 22
   6.1 Participants ......................................................................................................................................... 22
   6.2 Study Questionnaire ............................................................................................................................ 23
   6.3 Hospital contacts ................................................................................................................................. 24
   6.4 PFDI-20 Questionnaire ....................................................................................................................... 24
   6.5 Clinical examination ............................................................................................................................ 25
   6.6 Bladder Diary ..................................................................................................................................... 26
   6.7 Cervical pathology ............................................................................................................................... 26
   6.8 Ethical considerations .......................................................................................................................... 27
   6.9 Statistical analyses ............................................................................................................................... 28
   6.10 Meta-analyses of long-term follow-ups .............................................................................................. 29
7. Results ...................................................................................................................................................... 30
   7.1 Participants in the RCT .......................................................................................................................... 30
   7.2 Lower Urinary tract function .............................................................................................................. 33
7.2.1 Urinary incontinence ................................................................. 33
7.2.2 Other lower urinary tract symptoms ............................................. 36
7.2.3 Objective assessment of lower urinary tract function ...................... 38
7.3 Pelvic organ prolapse ................................................................. 40
7.4 Pain, constipation and hospital contacts ........................................... 41
7.5 Satisfaction with sex life ............................................................... 41
7.6 Quality of life ............................................................................. 44
7.7 Cervical/vaginal screening and pathology ......................................... 44
7.8 Meta-analyses of long-term follow-ups ............................................. 48

8. Discussion ................................................................................... 53
8.1 Overall findings .......................................................................... 53
8.2 Strengths .................................................................................. 53
  8.2.1 Randomized clinical trial ........................................................... 54
  8.2.2 Long-term follow-up ................................................................. 54
  8.2.3 Sample size ........................................................................... 54
  8.2.4 Cervical screening/pathology .................................................... 55
8.3 Limitations ................................................................................ 55
  8.3.1 Blinding ............................................................................... 55
  8.3.2 Missing data .......................................................................... 56
  8.3.3 Mass significance .................................................................. 59
  8.3.4 Sample size .......................................................................... 59
8.4 Interpretation ............................................................................. 60
  8.4.1 Urinary incontinence ............................................................... 60
  8.4.2 Urinary Flow ......................................................................... 62
  8.4.3 Pelvic organ prolapse ............................................................... 63
  8.4.4 Cervical cancer screening ........................................................ 64

9. Conclusion .................................................................................. 66
10. Recommendations ....................................................................... 66
11. Future aspects ........................................................................... 67
12. References ................................................................................ 69
13. Appendices ...................................................................................................................... 76
13.1 Questionnaire SAH ........................................................................................................ 76
13.2 Questionnaire TAH ....................................................................................................... 76
13.3 Questionnaire PFDI-20 ............................................................................................... 76
13.4 Bladder Diary .............................................................................................................. 76
14. Articles ........................................................................................................................ 76
**Abbreviations**

SAH: Subtotal abdominal hysterectomy

TAH: Total abdominal hysterectomy

UI: Urinary Incontinence

RCT: Randomized clinical trial

OS: Observational Study

MI: Multiple Imputation

ITT: Intention-to-treat

PP: Per Protocol

Qmax: Maximum flow rate in urinary flow

POP: Pelvic organ prolapse

POP-Q: Pelvic organ prolapse quantification system

QoL: Quality of Life

SF-36: Short Form 36, validated quality of life questionnaire

RR: Relative Risk

CI: Confidence Interval

PFDI-20: Pelvic Floor Distress Inventory 20

AJOG: American Journal of Obstetrics and Gynecology

IUJ: International Urogynecology Journal

EJOG: European Journal of Obstetrics and Gynecology and Reproductive Biology

DMJ: Danish Medical Journal
Preface

This research was carried out between 1st of August 2012 and 1st of July 2015. I was employed as a research fellow at the department of obstetrics and gynaecology, Nykøbing Falster Hospital and enrolled as a PhD-student at The University of Southern Denmark, Institute for Regional Health Research. The study was conducted in collaboration with the department of obstetrics and gynaecology, Roskilde Hospital and Juliane Marie Centret, Rigshospitalet.

I would like to express my huge gratitude towards my main supervisor Helga Gimbel MD, DMSc for trusting me with the follow-up of her randomized clinical trial which has been a part of her life for the past 20 years. Helga Gimbel has been supportive, helpful and available whenever I needed help. I have enjoyed all our meetings, methodological and clinical discussions and also our travels to conferences and meetings throughout the PhD.

Furthermore, I would like to thank my co-supervisors Professor Bent Ottesen MD, DMSc, and head of department Lars Alling Møller MD, PhD for their support and trust in me. I would like to thank them for their help with the initiation of the follow-up, critical reviews of the manuscripts and for practical help with equipment, office space, examination rooms and more. In addition, I would like to thank the entire original “Danish hysterectomy trial group” who put a huge effort into carrying out the randomized clinical trial and helped review and revise manuscripts for this thesis.

I would like to express my gratitude to nurses Inger Mathiasen and Anne Mette Karlsen for their professional help with the clinical examinations. I could not have done it without you.

Thank you to all my colleagues at the JMC research unit for women and children’s health and especially those I have shared an office with over the years: Hilde Hylland, Karina Jordan, Signe
Hanghøj, Majbritt Norman Nielsen, Agnethe Vale Nielsen and Mette Calundann Noer for support, laughs, academic discussions during lunch break and “friendly ears” when needed.

Further, I would like to thank Dr. Elizabeth Mueller, Dr. Linda Brubaker and the rest of the team at the department of “female pelvic floor medicine and reconstructive surgery”, Loyola University Hospital, Chicago, Illinois for allowing me to visit your department, be inspired by your work and discuss my research with you.

A special thanks to my beloved “better half” Kasper Munck, Business Delivery Manager, SAS institute for his great help with statistics and technical issues, for helping me with multiple imputation, for statistical discussions and for his support of me throughout the PhD. I would also like to thank my mother Maggie Laird Andersen for English proof reading of my articles and the thesis; I hope no commas are out of place.

Thank you to my friend Laura Franklin for designing the cover of the thesis and a special thanks to my two children Petra and Sixtus for their hugs, kisses and firm belief that Mum knows everything; after all she is a mother and a doctor — I am enjoying it while it lasts.

Last but not least I would like to thank all the women who participated in the randomized clinical trial and observational study and who agreed to participate in this long-term follow-up. I would like to dedicate this thesis to you as none of it would have been possible without your participation and interest in the trial.

The PhD was supported financially by the Research Unit in Region Sjælland, The University of Southern Denmark, Nykøbing Falster Hospital, Roskilde Hospital, Rigshospitalet, the Research Foundation in Region Sjælland, and “Lægernes Forsikringsforening af 1891.”

Copenhagen, June 2015

Lea Laird Andersen
1. Aim

The aim of this thesis was to conduct a follow-up to compare long-term consequences of subtotal and total abdominal hysterectomy in women who are now postmenopausal. The aim was to include:

- Subjective symptoms assessed by questionnaire: Urinary incontinence (primary outcome), constipation, quality of life, pelvic organ prolapse, pain, satisfaction with sex-life, hospital contacts, vaginal bleeding after SAH (secondary outcomes), and lower urinary tract symptoms (exploratory analyses).

- Objective outcomes assessed at clinical examination: Pad-weighing test, urinary flow, residual urine, POP-Q measurement, and bladder diary.

- Cervical cancer screening and cervical/vaginal pathology post-hysterectomy based on registry data.
The thesis is based on the following four articles described in table 1:

**Article 1:** Subtotal versus total abdominal hysterectomy: randomized clinical trial with 14-year questionnaire follow-up (Published in AJOG).¹

**Article 2:** Lower urinary tract symptoms after subtotal versus total abdominal hysterectomy: Exploratory analyses from a randomized clinical trial with a 14-year follow-up (Published in IUJ).

**Article 3:** Objective comparison of pelvic organ prolapse and urinary incontinence after subtotal vs. total abdominal hysterectomy: a randomized controlled trial with 14-year follow-up (Published in EJOG).

**Article (manuscript) 4:** Cervical/vaginal cytology after subtotal and total abdominal hysterectomy: 14 year follow-up of a randomized clinical trial as well as an observational study (Accepted on 05th of October, 2015 for publication in DMJ).
Table 1: Overview of the four articles this thesis is based on.

<table>
<thead>
<tr>
<th>Manuscript</th>
<th>Population</th>
<th>Participants</th>
<th>Data</th>
<th>Outcomes</th>
<th>Statistics</th>
<th>Results</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Subtotal versus total abdominal hysterectomy: randomized clinical trial with 14-year questionnaire follow-up (Published in AJOG)</td>
<td>RCT, 1996-2000: Hysterectomy on benign indication N=319 (SAH 161, TAH 158)</td>
<td>10 had died; 9 left Denmark 904 invited 907 (64.8%) participated (SAH 97, TAH 100) Follow-up time 14.1 years, mean age at follow-up 60.1 years</td>
<td>Study Questionnaire validated</td>
<td>Primary: UI (often or always), secondary: POP, pain, constipation, QoL, satisfaction with sex-life, hospital contacts and vaginal bleeding</td>
<td>X²-test, test and Wilcoxon’s rank sum test as appropriate, multiple imputation analysis for missing data. Intention-to-treat</td>
<td>Primary: 32 (33.3%) urinary incontinent after SAH, 20 (20%) after TAH (P=0.035). No difference in secondary outcomes. No difference between SAH and TAH in the multiple imputation analyses.</td>
<td>More women have subjective UI after SAH than after TAH. No difference in secondary outcomes. Interpret cautiously because of missing data and no significant difference in the multiple imputation model.</td>
</tr>
<tr>
<td>2. Lower urinary tract symptoms after subtotal versus total abdominal hysterectomy: exploratory analyses of a randomized clinical trial with a 14-year follow-up (accepted by IUJ)</td>
<td>RCT, 1996-2000: Hysterectomy on benign indication N=319 (SAH 161, TAH 158)</td>
<td>10 had died; 9 left Denmark 904 invited 907 (64.8%) participated (SAH 97, TAH 100) Follow-up time 14.1 years, mean age at follow-up 60.1 years</td>
<td>Study Questionnaire validated</td>
<td>Exploratory analyses: LUTS, UI subtypes, multivariate analyses of factors associated with UI and LUTS</td>
<td>X²-test, Wilcoxon’s rank sum test for non-normally distributed continuous data, multivariate logistic regression for multivariate analysis. Intention-to-treat</td>
<td>Stress UI was seen more often after SAH (N=40, 41.7%) than after TAH (N=27, 27%) RR: 1.54 95% CI: 1.03-2.3, P=0.03 No difference in other LUTS. Preoperative UI, local estrogen treatment and BMI≥25kg/m² are associated with UI 14 years after hysterectomy.</td>
<td>The difference in UI is mainly seen in SUI. High BMI and preoperative UI are related to UI 14 years after hysterectomy. No difference between SAH and TAH regarding any other LUTS.</td>
</tr>
<tr>
<td>3. Objective comparison of pelvic organ prolapse and urinary incontinence after subtotal vs. total abdominal hysterectomy: a randomized controlled trial with 14-year follow-up (Submitted to EJOG)</td>
<td>RCT, 1996-2000: Hysterectomy on benign indication N=319 (SAH 161, TAH 158)</td>
<td>100 (32.9%) participate in clinical examinations SAH 53, TAH 47 Follow-up time 14.1 years, mean age at follow-up 60.1 years</td>
<td>Objective measurements and PFDI-20 questionnaire</td>
<td>POP (POP-Q) UI (pad-weighing test, voiding function (urinary flow and bladder diary) bother related to the pelvic floor (PFDI-20 questionnaire)</td>
<td>X²-test for categorical outcomes, Wilcoxon’s rank sum test for non-normally distributed continuous data. Intention-to-treat</td>
<td>After SAH 31 (59.6%) women had objective stage 2 POP compared with 33 (70.2%) after TAH (P=0.27). Higher median (Qmax) after SAH than after TAH (P= 0.042). Higher median functional bladder capacity after SAH than after TAH (P=0.0147) according to bladder diary.</td>
<td>No difference in overall objective POP or UI. Higher Qmax and functional bladder capacity after SAH – could indicate larger bladder capacity. Risk of bias and insufficient power because of low participation.</td>
</tr>
<tr>
<td>4. Cervical/vaginal cytology after subtotal and total abdominal hysterectomy: 14 year follow-up of a randomized clinical trial as well as an observational study. (Submitted to DMJ)</td>
<td>RCT, 1996-2000 N=319 and observational study OS) (N=185) in the same period</td>
<td>801 women (RCT: 117, OS:184. SAH: 259, TAH: 242). Mean follow-up time 4.1 years and mean age at follow-up: 62.1 years. Women were followed from hysterectomy until 1) death, 2) migration, 3) 55th birthday or 4) 1st of Feb. 2014</td>
<td>Registry data from Patobank, Invitations to and withdrawals from screening, adherence to screening and pathological smears</td>
<td>Descriptive study X²-tests for categorical outcomes and -test for continuous outcomes. As treated</td>
<td>Adherence to screening after SAH was 61.4 %. After SAH 9.7% were not invited to screening. 8.5% had no smears done since hysterectomy. After TAH 34.3% had at least one smear taken since hysterectomy. 28 from the SAH group and one from the TAH group had at least one abnormal smear. No cervical cancers were found.</td>
<td>One in ten was excluded from the screening program after SAH and could be at increased risk of cervical cancer. One in three women had unnecessary tests done after TAH. A guideline on post-hysterectomy cervical/vaginal screening is warranted in order to target the correct women for screening.</td>
<td>No difference in overall objective POP or UI. Higher Qmax and functional bladder capacity after SAH – could indicate larger bladder capacity. Risk of bias and insufficient power because of low participation.</td>
</tr>
</tbody>
</table>
2. Summaries

2.1 English summary

This thesis is based on 4 articles (Articles 1-4). The aim was to compare subtotal (SAH) and total (TAH) abdominal hysterectomy for benign indications 14 years after the operation.

The follow-up is based on a randomized clinical trial (RCT) carried out in 1996 to 2000. An observational study (OS) carried out simultaneously with the RCT was also included in a registry-based study of cervical cancer screening after hysterectomy.

The women who participated in the RCT were followed-up by a questionnaire covering the pre-specified outcomes: Urinary incontinence (primary outcome), constipation, quality of life, pelvic organ prolapse, pain, satisfaction with sex-life, hospital contacts, vaginal bleeding after SAH (secondary outcomes) and lower urinary tract symptoms (exploratory analyses).

An objective assessment of lower urinary tract function and pelvic organ prolapse including pad-weighing test, urinary flow, residual urine, POP-Q measurement, and bladder diary was performed.

Finally assessment of the cervical cancer screening program was carried out. Data were retrieved from Patobank regarding adherence to the cervical cancer screening program and cervical pathology after hysterectomy.

We contacted 304 women. The Questionnaire was answered by 197 (64.8%) women (SAH: 97, TAH: 100). Mean follow-up time was 14.1 years and age at follow up was 60.1 years.

Urinary incontinence was more common after SAH (n= 32, 33.3%) than after TAH (n=20, 20%) (RR: 1.67, 95%CI: 1.02-2.70, P=0.035) (Article 1). Exploratory analyses showed that this difference was
primarily due to a difference in the number of women with stress urinary incontinence after SAH and TAH (*Article 2*). No difference was found between SAH and TAH regarding the secondary outcomes. Of those who answered the questionnaire, 100 (SAH: 53, TAH: 47) women participated in clinical examinations (*Article 3*). After SAH 31 (59.6%) women had objective pelvic organ prolapse compared with 33 (70.2%) after TAH (P=0.27). We found a higher median maximum flow rate (Qmax) after SAH (30.3 ml/sec) than after TAH (27.1 ml/sec) (P= 0.042) and a higher median functional bladder capacity after SAH (531.7ml) than after TAH (443.3 ml) (P=0.0147) according to the bladder diary. The registry study of cervical screening (*Article 4*) included 501 women (SAH: N= 259, TAH: N= 242). Almost one in ten (9.7%) of the women were not invited to screening after SAH. Adherence to screening was 61.4%; 30.1% were screened less frequently than every 5 years and 8.5% were not screened at all. After TAH 14.5% were not invited, 6.6% adhered to screening and 65.7% were not screened. We found minimum one abnormal test in 28 (10.8%) women after SAH and one after TAH. No cervical cancers were found.

In conclusion, we found no evidence favouring SAH compared with TAH; contrarily we found more cases of urinary incontinence after SAH. Furthermore, we found that cervical cancer screening after hysterectomy is not optimal: Some SAH women were not invited and most TAH women were invited, though they did not have a cervix. One in ten had abnormal cervical smears after SAH. The long-term outcomes presented in this thesis suggest that total hysterectomy should be the first choice of hysterectomy for benign indications.
2.2 Dansk Resumé

Afhandlingen er baseret på fire artikler (Artikel 1-4). Formålet med afhandlingen var at sammenligne subtotal (SAH) og total (TAH) abdominal hysterektomi på benign indikation med hensyn til langtidsfølger 14 år efter operationen.


Kvinderne modtog et spørgeskema der omhandlede urininkontinens (det primære effektmål), prolap, smerter, obstipation, tilfredshed med seksuallivet, livskvalitet, hospitalskontakter og vaginal blødning efter SAH (sekundære effektmål) samt urininkontinens subtyper og andre vandladningsgener (udybende analyser).

En objektiv klinisk undersøgelse bestående af blejevingstest, uroflowmetry, residualurinmåling, POP-Q måling og væskevandladningsskema blev også foretaget.

Heredover foretog vi en evaluering af cervix cancer screening efter hysterektomi på baggrund af kvinderne i det randomiserede studie samt det sideløbende observationsstudie. Data om deltagelse i screening samt patologisvar blev indhentet fra Patobank.

Vi kontaktede 304 kvinder. Spørgeskemaet blev besvaret af 197 (64,8 %) kvinder (SAH: 97, TAH: 100) (Artikel 1). Gennemsnits opfølgningstid var 14,1 år og gennemsnitsalderen ved opfølgningen var 60,1 år. Antallet af kvinder der angav at være urininkontinente var 32 (33,3 %) efter SAH og 20 (20 %) efter TAH (RR: 1,67, 95 % CI: 1,02-2,70, P=0,035). Uddybende analyser viste at denne forskel primært skyldtes forskel i hyppigheden af stress urininkontinens mellem SAH og TAH (Artikel 2). SAH og TAH var sammenlignelige med hensyn til de sekundære effektmål.
Af dem der besvarede spørgeskemaet deltog 100 (SAH:53, TAH: 47) kvinder i kliniske undersøgelser (Artikel 3). Efter SAH observeredes grad 2 prolaps hos 31 (59,6 %) kvinder sammenlignet med 33 (70,2 %) kvinder efter TAH (P=0,27). Vi fandt højere median maksimal flow rate (Qmax) hos kvinder efter SAH (30,3 ml/sec) end hos kvinder efter TAH (27,1 ml/sec) (P= 0,042) og højere median funktionel blærekapacitet efter SAH (531.7ml) end efter TAH (443.3 ml) (P=0,0147) ifølge væskevandladningsskemaet.

Studiet om cervix cancer screening inkluderede 501 kvinder (SAH: 259, TAH: 242) (Artikel 4). Efter SAH blev 9,7 % af kvinderne ikke inviteret til screening. Tilslutning til cervix cancer screening efter SAH var 61,4 %; 30,1 % blev screenet mindre end hver femte år og 8,5 % blev ikke screenet. Efter TAH blev 14,5 % ikke inviteret til screening, tilslutning til screening var 6,6 % og 65,7 % blev ikke screenet. Hos 28 (10,8 %) kvinder i SAH gruppen fandt vi mindst én abnorm prøve. I TAH gruppen havde én kvinde en abnorm prøve. Alle abnorme prøver blev fulgt op med relevante undersøgelser. Vi fandt ingen tilfælde af cervix cancer.

Vi har ikke fundet evidens for at SAH er TAH overlegen. Tværtimod, fandt vi flere urininkontinente kvinder i gruppen der fik foretaget SAH. Ydermere, fandt vi at screening for cervix cancer efter hysterektomi ikke er optimal: Nogle kvinder blev ikke inviteret efter SAH og de fleste kvinder blev inviteret efter TAH selvom de ikke har en cervix. Én ud af ti kvinder oplevede abnorme celleprøver efter SAH. Ifølge resultaterne fra denne afhandling bør total hysterektomi være førstevalg ved benign sygdom.
3. Background

3.1 History:

Hysterectomy dates back to the ancient Greeks who performed vaginal hysterectomies for prolapsed uteri.\(^2\) The abdominal hysterectomy was first attempted in the middle of the 19\(^{th}\) century with mortality rates of more than 70\%.\(^2\) The discussion of subtotal versus total hysterectomy has been ongoing for decades. The subtotal hysterectomy was the first choice before the introduction of safe anaesthesia and the discovery of antibiotics\(^2\) as lengthy anesthesia and contamination from the vagina were more likely to be fatal than a potential cervical cancer. After antibiotics were introduced, the total abdominal hysterectomy became the first choice, when possible, to avoid future cervical cancers in these women.\(^2\) In the 1980s a new shift towards the subtotal procedure was seen because studies indicated that the subtotal procedure was of benefit regarding women’s sexuality.\(^3,4\) In recent years the mode of hysterectomy has changed. A shift towards less invasive methods such as laparoscopic hysterectomy is seen\(^5\) and the issue of whether the cervix should be removed continues: The surgeons found the subtotal method easier to perform; however, the subtotal laparoscopic hysterectomy method is now discouraged because of the risk of spread of leiomyosarcoma due to morcellation which is a part of the procedure.\(^6\)

3.2 Hysterectomy in Denmark

Approximately 4500 hysterectomies are performed for benign indications each year in Denmark.\(^7,8\) Even though other less invasive methods have been developed for treating bleeding disorders and fibromas,\(^5\) the rate of hysterectomy has remained stable at about 180 per 100,000 women years.\(^5\)
In recent years the abdominal hysterectomy has given way to the less invasive laparoscopic hysterectomy which has been found to be superior regarding hospital stay, return to normal activities, blood loss and wound infections. The percentage of hysterectomies performed as vaginal hysterectomy has decreased from 32% in 2011 to 22% in 2013 despite the fact that the vaginal approach was found to be comparable to the laparoscopic approach regarding most outcomes and actually favourable regarding surgical time. In 2013 27% of the hysterectomies in Denmark were still abdominal. This has decreased from 39% in 2011, whereas the percentage of hysterectomies performed laparoscopically has increased from 29% in 2011 to 51% in 2013.

Of all benign hysterectomies in Denmark in 2011 8.8% were subtotal: 10.4% of abdominal and 19.2% of laparoscopic hysterectomies were subtotal. All vaginal hysterectomies were total. In Denmark the percentage of hysterectomies performed as subtotal hysterectomies varies between departments from 0-39% and in some departments in Germany subtotal laparoscopic hysterectomy is the standard offered to all women without contraindications. A recent national guideline regarding hysterectomy on benign indications recommends that vaginal hysterectomy should be considered the first choice if there are no contraindications. Furthermore it states that minimally invasive methods are to be used whenever possible.

### 3.3 Subtotal versus total hysterectomy

Supporters of subtotal hysterectomy have argued that shorter surgical time and quicker recovery, as well as a less invasive surgery would result in better uro-genital outcomes because of fewer adhesions and less disruption of the nerve innervations of the pelvic organs. A recent non-randomized trial comparing subtotal with total laparoscopic hysterectomy found that women who had undergone subtotal hysterectomy resumed normal activities faster than women who had undergone total...
hysterectomy. The cervix’ role in sexuality and orgasm has been discussed. Prospective cohort studies in the 1980s suggested that SAH was of benefit to the women in sexual matters. These findings have not been reproduced in RCTs nor in the systematic review. One RCT has shown a significantly better body image one year after subtotal compared with total abdominal hysterectomy. Another study found that women who underwent subtotal hysterectomy resumed sexual activity faster, had higher sexual desire and more stated that sexual life had improved. However, this was not a randomized trial and could be prone to selection and recall bias.

Supporting the total hysterectomy a Cochrane review concluded that even though the surgical time was shorter for subtotal hysterectomy this was unlikely to have a clinical benefit and they found no difference in recovery time based on hospital stay and return to normal daily activities. The same review found no difference in urinary or bowel function to support that subtotal hysterectomy would lead to better uro-genital outcomes. The RCT this thesis is based on found that one year after SAH 24 (18%) had urinary incontinence compared with 13 (9%) after TAH, RR: 1.89, 95%CI: 1.004-3.55, P= 0.043. At the five-year follow-up the difference between SAH and TAH in urinary incontinence persisted (SAH: 34 (30.1%), TAH: 21 (17.6%), RR: 1.7, 95%CI 1.06-2.75, P= 0.026).

A disadvantage of the subtotal hysterectomy procedure is continued cyclical bleeding or spotting. We have previously shown that 20% of the women in the SAH group had vaginal bleeding or spotting at one year and after 5 years 10.9% of the women in the SAH group still experienced vaginal bleeding or spotting.
Morcellation\textsuperscript{23-26} which is a part of the subtotal laparoscopic hysterectomy procedure is associated with a risk of disseminating fibromas or spreading undiagnosed leyomyosarcomas. Because of this morcellation is no longer recommended by the food and drug administration (FDA) in the US.\textsuperscript{6}

The most recent Cochrane review\textsuperscript{19} comparing subtotal and total hysterectomy included both open and laparoscopic surgery and stated that no evidence supports favouring subtotal hysterectomy in laparoscopic or abdominal surgery. The review\textsuperscript{19} concludes that more long-term follow-up is needed, as uro-genital problems in particular, may occur years after surgery — especially in postmenopausal women.

The recent Danish national guideline recommends against subtotal hysterectomy because of the risk of continued vaginal bleeding and the need for cervical cancer screening associated with the subtotal procedure.\textsuperscript{12}

### 3.4 Cervical cancer screening

A reason for opposing subtotal hysterectomy has been the risk of developing cervical cancer in the remaining cervix and thereby the continued need for cervical cancer screening.\textsuperscript{13} Screening for a disease is recommended if 1) the disease is serious 2) the disease is preventable or treatable when found at an early stage, 3) a safe and valid test causing no harm and only little inconvenience is available.

To prevent cervical cancer a national cervical screening program as well as an HPV vaccine is offered to women in Denmark. Since 1967 regional screening programs have been available in Denmark. In 1996 a national cervical cancer screening program was implemented.\textsuperscript{27} In the current form women are
invited for screening every 3 years from the age of 23 till 50 and every 5 years from 50-65. From the age of 23-59 the screening is cytology based whereas an HPV test has recently been implemented as the primary screening of women aged 60-64. The screening programs are estimated to have prevented almost half the cases of cervical cancer that would have been expected without screening in the Nordic countries.\textsuperscript{27} The HPV vaccine was implemented as part of the Danish children’s vaccination program in 2008 and is offered to girls at the age of 12. It is estimated that it will protect against the 70\% of cervical cancers caused by the types of HPV covered by the vaccine.\textsuperscript{28} A guideline regarding cervical cancer screening in relation to hysterectomy is not available in Denmark. If a cervical smear is done prior to surgery showing no abnormal cells the risk of post hysterectomy cervical cancer should be minimal.\textsuperscript{13} Storm et al.\textsuperscript{29} found no overall increased risk of cervical cancer or other cancers among women who had undergone subtotal hysterectomy compared with the general population. However, they did find an increased risk of cervical cancer among women who were above the age of 50 at time of subtotal hysterectomy. A study by Hellstrom et al\textsuperscript{30} found that cervical stump cancers account for 2\% of all cervical cancers and that average time from hysterectomy until diagnosis is 17.6 years. They found that patients with cervical stump cancers did not have a worse stage-by-stage prognosis compared with cervical cancer patients with intact uteri. However, as the stump cancers tended to be diagnosed at a later stage they had an overall worse prognosis. They suggest that insufficient screening of women after subtotal hysterectomy might be responsible for the late diagnosis resulting in a worse prognosis.\textsuperscript{30} Smears of the vaginal vault are of no proven value after total hysterectomy for benign indications, contrarily they can lead to anxiety and discomfort for the patient and unnecessary costs.\textsuperscript{31} In the UK there are national guidelines as to whom should have vaginal vault smears performed after total hysterectomy.\textsuperscript{32} In the US the preventive services task force recommends against screening after total
hysterectomy for benign indications.³³ A guideline regarding post-hysterectomy cervical/vaginal screening does not exist at present in Denmark.

4. **Aim**

The aim of this thesis was to conduct a follow-up to compare long-term consequences of subtotal and total abdominal hysterectomy in women who are now postmenopausal. The aim was to include:

- **subjective symptoms assessed by questionnaire:** Urinary incontinence (primary outcome), constipation, quality of life, pelvic organ prolapse, pain, satisfaction with sex-life, hospital contacts, vaginal bleeding after SAH (secondary outcomes) and lower urinary tract symptoms (exploratory analyses).

- **Objective outcomes assessed at clinical examination:** Pad-weighing test, urinary flow, residual urine, POP-Q measurement, and bladder diary.

- **Cervical cancer screening and cervical/vaginal pathology post-hysterectomy based on registry data**
5. Setting

The comparison was based on women included in a Danish multi-centre RCT 14 years ago, comparing subtotal with total abdominal hysterectomy for benign diseases (except pelvic organ prolapse). For the study of cervical cancer screening and cervical/vaginal pathology, women from the RCT as well as a simultaneously conducted observational study (OS) were included. Data regarding hospital contacts since hysterectomy and cervical cancer screening and pathology were obtained from registries (National Patients’ registry and Patobank, respectively)

6. Methods

6.1 Participants

Women referred for hysterectomy on benign indication were recruited to a randomized clinical trial (RCT) at 11 gynaecological departments from 1996 to 2000 in Denmark. The women were randomized to subtotal abdominal hysterectomy or total abdominal hysterectomy. No instructions were given to the surgeon except to electrocoagulate the cervical canal during the subtotal procedure. The women in the trial were not blinded to the treatment. The sample size was based on a prevalence of urinary incontinence after TAH of approximately 23%. With a power of 0.80, a type I error of 5%, and a 15% absolute difference in urinary incontinence between the two operations, 160 participants were needed in each intervention group. In- and exclusion criteria have been published.

Data from questionnaires prior to surgery, after 2 months, 6 months, one and 5 years have been published previously.
For this long-term follow-up all women included in the original trial who were still alive and living in Denmark were contacted by letter in October 2012 and asked to answer the study questionnaire as in previous follow-ups (*Appendix sections 13.1, 13.2*) (Figure 1). The letter included information about the long-term follow-up and an invitation to participate in clinical examinations. Reminders were sent to non-responders after two and 7 months. Alongside inclusion to the RCT an observational study (OS) was carried out. The women from the OS were included in this follow-up regarding cervical cancer screening and pathology (Figure 6). The exclusion criteria for the OS and the RCT were the same. Women undergoing hysterectomy for pelvic organ prolapse as well as women with any previous abnormal smears were excluded. All women had a normal smear prior to hysterectomy.

### 6.2 Study Questionnaire

Outcomes regarding subjective symptoms of urinary incontinence (often or always, primary outcome), pelvic organ prolapse, constipation, pain, sexuality, quality of life (SF-36) and vaginal bleeding after SAH (secondary outcomes) as well as lower urinary tract symptoms and urinary incontinence subtypes: “in which situations do you experience urinary leakage?” (Exploratory outcomes) were obtained from the study questionnaire developed for the trial. The questionnaire was validated prior to the initiation of the RCT. The validation included content validity and a patient-interviewer agreement study. Some changes were made to the questionnaire based on the validity study before it was used in the RCT. Details of the scoring and assessment of the different outcomes are described in *Article 1* and *Article 2*. 

23
6.3 Hospital contacts

All hospital contacts derived from discharge summaries of hospitalizations or outpatient treatment from Danish public hospitals were registered from time of surgery till July 2013. Any hospital contact regarding abdominal, gynaecological, urological (including urinary incontinence), plastic surgical, or dermatological complaints were scrutinized. Contacts possibly related to the prior hysterectomy were registered (Article 1). Hospital contacts and complications to the hysterectomy from time of surgery till 5 years postoperatively have been published previously but are included in the results section to show the full picture.

6.4 PFDI-20 Questionnaire

To obtain further information about bother related to pelvic floor symptoms we used the Pelvic Floor Distress Inventory questionnaire (PFDI-20, Appendix section 13.3) in addition to the study questionnaire (Article 3). The PFDI-20 was filled-in by the women who had stated on the main study questionnaire that we could contact them with further questionnaires. The PFDI-20 has been validated in Danish. The PFDI-20 score was obtained from the three subscales: Pelvic Organ Prolapse Distress Inventory (POPDI-6), Colorectal-Anal Distress Inventory (CRAD-8) and Urinary Distress Inventory (UDI-6). The score was calculated by multiplying the mean score (0-4) in each subscale by 25 and adding them up, resulting in three sub scores between 0 and 100 and a PFDI-20 score between 0 and 300. Higher scores indicate more bothersome pelvic floor symptoms.
6.5 Clinical examination

We invited all women from the RCT to participate in clinical examinations (Article 3). The examinations consisted of:

- 20-minute pad-weighing test as specified by Machold et al\textsuperscript{40}
- Urinary flow
- Measurement of residual urine (by ultrasound)
- Gynaecological examination including POP-Q measurement
- Cervical smear in the SAH group

SAH women from the OS were also invited for the cervical smear unless they stated they had a normal smear result from within the last year prior to examination.

Participants were asked not to void for two hours prior to the appointment aiming at a comfortably full bladder (about 200ml) at time of pad-weighing test and urinary flow. POP-Q measurement for pelvic organ prolapse was carried out in a supine position after urinary flow, and measured as specified by Bump et al.\textsuperscript{41} Methods, definitions, and descriptions conform to the standards recommended by the International Continence Society (ICS).\textsuperscript{41} The woman did the Valsalva maneuver; measurements were done during maximum pressure. Stage 2 or more, according to POP-Q measurement,\textsuperscript{42} was defined as POP. Staging was based on the most distal measurement and categorized according to in which compartment this was noted: Anterior (point Aa or Ba), apical (point C or D) or posterior (Point Ap or Bp). If the anterior and posterior protrusion were the same it was categorized as “Both”.

The flowmeter measures voided volume and time and draws a curve of the voiding function with time on the x-axis and flow rate (ml/seconds) on the y-axis.\textsuperscript{43} Qmax and voided volume were registered at
urinary flow and compared with the expected values according to the Liverpool Nomogram using the formula for the female nomogram: \( \ln(Q_{\text{max}}) = 0.511 + 0.505 \times \ln(\text{Voided volume}) \). A normal flow is seen as a smooth curve indicating no straining and no obstruction.\(^{44}\) A normal curve indicates normal voiding function whereas abnormal urinary flow can occur for several reasons. Another measurement of the emptying function of the bladder is the residual urine. We measured residual urine by ultrasound to assess whether the bladder was sufficiently emptied. A specific cut-off for abnormal residual urine is not clearly defined; however, a cut-off of 50-100ml is suggested by IUGA/ICS.\(^{45}\) We chose the cut-off of 100ml.

### 6.6 Bladder Diary

Following the clinical examination the women kept a bladder diary (Appendix section 13.4) for three consecutive days and nights (Article 3). They registered fluid intake, urinary output, use of sanitary pads and leakage of urine, and under which circumstances it occurred. No urinary leakage during the 3 day bladder diary was considered as continent. Nocturia was defined as waking up to void minimum once during the night (18). As no relevant international definition to our concern was available we defined urinary frequency as more than eight voiding episodes daily.

### 6.7 Cervical pathology

We obtained registry data from the national pathology databank, Patobank (Article 4). The data included all invitations to screening, reminders, active withdrawals by the women (and reasons) as well
as cervical/vaginal pathology results for all women included in the randomized clinical trial (RCT) and observational study (OS).

The individual follow-up time for each participant was from date of hysterectomy until the first of the following: Death, emigration, age 65 or 1st of February 2014. Age was included to focus on smears taken within the age group covered in the national screening program. Adherence to the cervical screening program was defined as minimum one test per five years of follow-up.

6.8 Ethical considerations

All participants signed an informed consent prior to the initiation of the trial.\textsuperscript{21} For this follow-up participants were informed by letter about the trial and about their rights as participants in biomedical research. The questionnaire took about 20-30 minutes to fill-in. All participants in clinical examinations signed an informed consent.

Gynaecological examinations are associated with some discomfort but do not involve health risks for the participants. We found it necessary to perform these examinations on all participants who agreed, including those without current symptoms, in order to generate comparable data for the two surgical groups.

The local ethics committees of the participating centers: Bornholm, Frederiksborg, Roskilde, Storstrøms, and Vestsjællands Counties, and the Danish Data Protection Agency had accepted the design of the randomized clinical trial before recruitment of patients. The Danish Data Protection Agency and the regional ethics committee in Region Sjælland accepted the design of the long term follow-up.
6.9 Statistical analyses

For categorical data $\chi^2$-tests were used. For continuous data we used t-tests for normally distributed data and Wilcoxon’s rank-sum test for non-normally distributed data. Analyses were performed as intention-to-treat. Non-responders were excluded (Articles 1-3). For the primary outcome (urinary incontinence, often or always) we chose to perform several additional analyses: 1) per protocol analysis excluding all women who did not have the type of hysterectomy they were randomized to, 2) analysis of the number of women who reported urinary incontinence at any follow-up since hysterectomy, 3) urinary incontinence according to study questionnaire or according to discharge summaries (Article 1). Furthermore, to account for missing data we performed multiple imputation analyses for the analyses of urinary incontinence, constipation, pelvic organ prolapse, satisfaction with sex life, pelvic pain, and quality of life (secondary outcomes). Conclusions are based on the regular intention-to-treat analyses. Multiple imputation (MI) was carried out using the FCS method in SAS® version 9.3 using the PROC MI and MIANALYZE functions. Multiple imputed data was analysed using Rubin’s rules, these incorporate within imputation and between imputation variance to account for the uncertainty of the imputed values. With high percentages of missing data this uncertainty increases; as a result the method is not valid with high percentages of missing data. The multiple imputation results should be compared with the results of analyses of observed data (complete case analyses); any differences should be addressed.

Multivariate logistic regression was carried out to consider associations between different variables and urinary incontinence and between LUTS and bothersome voiding (Article 2). Analyses of cervical cancer screening data (Article 4) were done according to “as-treated”.
All data were handled and analysed in SAS JMP® version 10 (SAS institute, Cary, North Carolina, USA) except MI as explained above. Statistical significance was defined as a two-sided P-value of less than 0.05.

6.10 Meta-analyses of long-term follow-ups

For this thesis a meta-analysis of trials comparing long-term outcomes after SAH and TAH was carried out. The trials included were limited to randomized clinical trials with follow-up time of more than 5 years. The intervention of interest was subtotal abdominal hysterectomy on benign indication and the comparison of interest was total abdominal hysterectomy on benign indication. Outcomes of interest were urinary incontinence including subtypes (Stress urinary incontinence and urgency urinary incontinence) and pelvic organ prolapse.

PUBMED and EMBASE as well as the Cochrane database were searched using the terms: Hysterectomy, subtotal, total, SAH, TAH, supravaginal, randomized clinical trials, follow-up, and long-term. In addition references in relevant articles were studied to find any additional studies.

Furthermore we included our own studies as presented in this thesis including those not yet published. The trials which were included were assessed regarding number of participants, blinding, follow-up time, missing data, and relevant outcomes. Data extraction and analyses were performed by the author of this thesis.

The meta-analyses were carried out in Review Manager version 5.3 using a fixed effects model which assumes that the true effect of the intervention is the same in all trials. This assumption was tested by $\chi^2$ and $I^2$ tests for heterogeneity. Odds ratios and 95% confidence intervals are presented.
7. Results

7.1 Participants in the RCT

We contacted 304 of the 319 women included in the RCT. Of the remaining, 10 had died of causes unrelated to hysterectomy and 5 had left Denmark. All the women were Caucasian. The study questionnaire was returned by 249 (82%) (SAH: 119/TAH: 130) women; however, 52 (17.1%) (SAH: 22/TAH: 30) returned it blank stating that they did not wish to participate; 197 (64.8%) answered the questionnaire (SAH: 97 (63.4%), TAH: 100 (66.2%)) (Article 1 and 2). Of those who answered the study questionnaire, 140 (46.1%) (SAH: 68, TAH: 72) answered the PFDI-20 questionnaire and 100 women (32.3 %) (SAH: 53 (34.6%) and TAH: 47(31.1%)) agreed to participate in clinical examinations (Article 3). The bladder diary was completed by 68/304(22.4%) (SAH: 34, TAH: 34) (Figure 1). Basic characteristics of the women did not differ significantly between surgical groups, responders and non-responders to the questionnaire or participants in clinical examinations except that fewer responders were smokers at time of surgery than non-responders and more responders had a high alcohol intake at time of surgery than non-responders (Table 2). Mean age of the women was 60.1 years and the mean follow-up time was 14.1 years.
Figure 1: Flow of the patients through each stage of the trial at 14 year follow up.

Randomised (n=319)

Allocation

Allocated to intervention (TAH): 158
- Received allocated intervention: 139
- Did not receive allocated intervention: 19
  Patient preference: 6
  Surgical/medical reasons: 12
  Not hysterectomized: 1

Allocated to intervention (SAH): 161
- Received allocated intervention n=138
- Did not receive allocated intervention: 23
  Patient preference: 15
  Surgical/medical reasons: 7
  Not hysterectomized: 1

Lost to follow-up:
1 year: 18
5 years: 39

Follow-Up

Lost to follow-up at 14 years: 58
- 3 had died
- 4 had left the country
- 30 did not wish to participate
- 21 did not respond

Lost to follow-up:
1 year: 25
5 years: 46

Analysis

Analysed (n=100) (ITT)

Excluded from analysis (n=13)
- Change of method: (n=12)
- Patient preference: 2
- Surgical/medical reasons: 10
- Did not meet inclusion criteria: (n=1) had continence surgery prior to hysterectomy

Analysed in per protocol: 87

Participated in clinical examination: 47

Filled out voiding diary: 34

Analysed (n=97) (ITT)

Excluded from analysis (n=13)
- Not hysterectomised: 1
- Change of method: (n=8)
- Patient preference: 4
- Surgical/medical reasons: 4
- Did not meet inclusion criteria: (n=4) 1 had diabetes mellitus, 2 had a diagnosis of malignant disease, 1 had alcohol problems

Analysed in Per Protocol: 84

Participated in clinical examination: 53

Filled out voiding diary: 34
Table 2 Characteristics of responders and non-responders to the study questionnaire and participants in clinical examinations.

<table>
<thead>
<tr>
<th></th>
<th>Responders</th>
<th>Non-Responders</th>
<th>P&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Participants in Clinical Examination</th>
<th>P&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All responders</td>
<td>SAH</td>
<td>TAH</td>
<td>N=122</td>
<td></td>
</tr>
<tr>
<td>Age mean (SD)</td>
<td>60.1 (5.8)</td>
<td>60.7 (5.9)</td>
<td>59.6 (5.6)</td>
<td>60.5 (6.6)</td>
<td>0.6</td>
</tr>
<tr>
<td>Follow-up time mean (range)</td>
<td>14.1 (12-16)</td>
<td>14.2 (12-16)</td>
<td>14.03 (12-16)</td>
<td>14.04 (12-16)</td>
<td>0.77</td>
</tr>
<tr>
<td>Parity mean (range)</td>
<td>1.8 (0-5)</td>
<td>1.85 (0-5)</td>
<td>1.76 (0-4)</td>
<td>1.74 (0-5)</td>
<td>0.6</td>
</tr>
<tr>
<td>BMI mean (SD)</td>
<td>26.1 (6.7)</td>
<td>26.45 (7.1)</td>
<td>25.71 (6.3)</td>
<td>25.5 (4.6)</td>
<td>0.38</td>
</tr>
<tr>
<td>Smoking &gt; 5 cigarettes per day&lt;sup&gt;d&lt;/sup&gt; n (%)</td>
<td>46 (23.4)</td>
<td>18 (18.6)</td>
<td>28 (28.0)</td>
<td>57 (46.7)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Alcohol &gt; 14 units per week&lt;sup&gt;e&lt;/sup&gt; n (%)</td>
<td>22 (11.2)</td>
<td>13 (13.4)</td>
<td>9 (9.0)</td>
<td>6 (4.9)</td>
<td>0.047</td>
</tr>
<tr>
<td>Urinary incontinence preoperatively</td>
<td>48/192 (25)</td>
<td>26 (27.7%) (n=94)</td>
<td>22 (22.2%) (n=99)</td>
<td>20/115 (17.4)</td>
<td>0.12</td>
</tr>
<tr>
<td>Indication for hysterectomy n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibroids</td>
<td>115 (58.4)</td>
<td>58 (59.8)</td>
<td>57 (57.0)</td>
<td>70 (57.4)</td>
<td>0.86</td>
</tr>
<tr>
<td>Dysfunctional uterine bleeding</td>
<td>63 (32.0)</td>
<td>29 (29.9)</td>
<td>34 (33.6)</td>
<td>42 (34.4)</td>
<td>0.65</td>
</tr>
<tr>
<td>Dysmenorrhea</td>
<td>8 (4.1)</td>
<td>3 (3.1)</td>
<td>5 (4.9)</td>
<td>4 (3.3)</td>
<td>0.72</td>
</tr>
<tr>
<td>Pelvic pain</td>
<td>9 (4.6)</td>
<td>6 (6.2)</td>
<td>3 (2.9)</td>
<td>4 (3.3)</td>
<td>0.57</td>
</tr>
<tr>
<td>Endometriosis</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1(0.8)</td>
<td>0.17</td>
</tr>
<tr>
<td>Other</td>
<td>2(1.0)</td>
<td>1(1.03)</td>
<td>1(0.99)</td>
<td>1(0.8)</td>
<td>0.86</td>
</tr>
<tr>
<td>Type of surgery n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAH</td>
<td>97 (49.2)</td>
<td>64 (52.5)</td>
<td>0.49</td>
<td>53 (53)</td>
<td>0.54</td>
</tr>
<tr>
<td>TAH</td>
<td>100 (51.8)</td>
<td>58 (47.5)</td>
<td>0.49</td>
<td>47 (47)</td>
<td>0.54</td>
</tr>
</tbody>
</table>

<sup>a</sup>Comparison of responders and non-responders to the questionnaire

<sup>b</sup>Comparison of those participating in clinical examination with all of those who did not.

<sup>c</sup>At 14 years for responders, preoperatively for non-responders

<sup>d</sup>At time of surgery

<sup>e</sup>At 14 year follow-up for responders and 1 year follow-up for non-responders
7.2 Lower Urinary tract function

7.2.1 Urinary incontinence

The analyses of the primary outcome of the 14 year follow-up — urinary incontinence (often or always) — showed significantly more women with urinary incontinence after SAH (n= 32, 33.3%) than after TAH (n= 20, 20%), RR: 1.67, 95% CI 1.02-2.70, P= 0.035, however the MI results showed no difference (Table 3). The additional analyses of urinary incontinence i.e. per protocol and urinary incontinence at any follow-up, or urinary incontinence according to the study questionnaire or discharge summary also showed more women with urinary incontinence after SAH than after TAH; however, not all showed statistically significant differences (Article 1).¹

The exploratory analyses of urinary incontinence subtypes (Article 2) showed that the difference found between SAH and TAH regarding urinary incontinence, is mainly due to a difference in the number of women experiencing leakage during sports/physical activity (i.e. stress urinary incontinence) (Table 3). The multivariate logistic regression analysis showed that being urinary incontinent prior to hysterectomy, BMI>25 kg/m² and local oestrogen treatment were significantly associated with being urinary incontinent 14 years after hysterectomy (Article 2). BMI>25 was mainly associated with urgency symptoms and mixed urinary incontinence (Table 3).

Figure 2 shows the percentage of women with urinary incontinence at each time point in the trial. The groups were comparable before hysterectomy. Consistent with the current follow-up there were significantly more urinary incontinent women in the SAH group than in the TAH group after one and five years. The percentage decreased in both groups after hysterectomy; however, more so in the TAH
group. Over the years the percentage increased in both groups; however more in the SAH group. In the TAH group the percentage at 14 years is comparable to the percentage prior to hysterectomy.
### Table 3: Urinary incontinence (UI) and UI subtypes according to type of hysterectomy and BMI

<table>
<thead>
<tr>
<th>Incontinence</th>
<th>SAH N=97</th>
<th>TAH N=100</th>
<th>RR</th>
<th>95% CI</th>
<th>P</th>
<th>BMI &gt;25kg/m² N= 108</th>
<th>BMI ≤25kg/m² N= 87</th>
<th>RR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Often or always (primary outcome) (N=97/100) [MI]</td>
<td>32 (33.3)</td>
<td>20 (20)</td>
<td>1.67</td>
<td>1.02-2.70</td>
<td>0.035*</td>
<td>36 (33.3)</td>
<td>16 (18.4)</td>
<td>1.8</td>
<td>1.08-3.04</td>
<td>0.019*</td>
</tr>
<tr>
<td>BMI ≤25kg/m²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI &gt;25kg/m²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In which situations do you experience urinary leakage?*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A:always</td>
<td>1(1.04)</td>
<td>1(1)</td>
<td>0.999</td>
<td>0.066-16.42</td>
<td>0.98</td>
<td>2(1.85)</td>
<td>0</td>
<td>0.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B:During intercourse</td>
<td>0</td>
<td>2(2)</td>
<td>0.099</td>
<td>0</td>
<td>0.00-2.83</td>
<td>0.013</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C: Urgent desire to void</td>
<td>31(32.3)</td>
<td>35(35)</td>
<td>0.92</td>
<td>0.62-1.37</td>
<td>0.69</td>
<td>46(42.6)</td>
<td>20(23)</td>
<td>1.85</td>
<td>1.19-2.89</td>
<td>0.004*</td>
</tr>
<tr>
<td>D:Cough, sneeze or laugh</td>
<td>46(47.9)</td>
<td>43(43)</td>
<td>1.11</td>
<td>0.82-1.52</td>
<td>0.49</td>
<td>55(50.9)</td>
<td>33(37.9)</td>
<td>1.34</td>
<td>0.97-1.86</td>
<td>0.069</td>
</tr>
<tr>
<td>E:Sports/physical activity</td>
<td>31(32.3)</td>
<td>16(16)</td>
<td>2.02</td>
<td>1.18-3.44</td>
<td>0.008*</td>
<td>25(23.2)</td>
<td>22(25.3)</td>
<td>0.92</td>
<td>0.56-1.51</td>
<td>0.73</td>
</tr>
<tr>
<td>F: Never</td>
<td>25(26.04)</td>
<td>37(37)</td>
<td>0.70</td>
<td>0.46-1.07</td>
<td>0.09</td>
<td>29(26.9)</td>
<td>33(37.9)</td>
<td>0.71</td>
<td>0.47-1.07</td>
<td>0.1</td>
</tr>
</tbody>
</table>

**Urinary incontinence (UI) subtypes**

<table>
<thead>
<tr>
<th>Mixed UI (C and D or E)</th>
<th>20(20.8)</th>
<th>18(18)</th>
<th>1.16</th>
<th>0.65-2.05</th>
<th>0.62</th>
<th>27(25.0)</th>
<th>11(12.6)</th>
<th>1.98</th>
<th>1.04-3.76</th>
<th>0.030*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urgency UI (C but no other)</td>
<td>11(11.5)</td>
<td>17(17)</td>
<td>0.67</td>
<td>0.33-1.36</td>
<td>0.27</td>
<td>19(17.6)</td>
<td>9(10.3)</td>
<td>1.7</td>
<td>0.81-3.6</td>
<td>0.15</td>
</tr>
<tr>
<td>Stress UI (D and/or E but no other)</td>
<td>40(41.7)</td>
<td>27(27)</td>
<td>1.54</td>
<td>1.03-2.3</td>
<td>0.03*</td>
<td>33(30.6)</td>
<td>33(37.9)</td>
<td>0.81</td>
<td>0.54-1.91</td>
<td>0.28</td>
</tr>
</tbody>
</table>

*Statistically significant P<0.05

bMultiple Imputation

*Question 34 in the postal questionnaire. The table consists of the possible answers (A-F), each woman can answer more than one category meaning that the total number in A-F is larger than the number of participants. It also includes women who only experience leakage rarely.
Figure 2: Percentage of urinary incontinent women in the SAH and TAH groups at each time point.

7.2.2 Other lower urinary tract symptoms

There was no difference between surgical groups regarding incomplete bladder emptying, straining to void, poor stream, frequency (>10 voids per day), nocturia, double/triple voiding, dysuria or urinary tract infection (Table 4). A multivariate analysis showed that urinary incontinence and sensation of incomplete emptying of the bladder were associated with bothersome voiding function (Table 4). There was no difference between SAH and TAH regarding urinary incontinence, pelvic organ prolapse or bowel symptoms as measured in the PFDI-20 scores (Article 3).
Table 4: Lower urinary tract symptoms according to hysterectomy groups and multivariate logistic regression of association of LUTS with bothersome voiding.

<table>
<thead>
<tr>
<th>Lower Urinary Tract symptoms</th>
<th>SAH</th>
<th>TAH</th>
<th>RR</th>
<th>95% CI</th>
<th>P</th>
<th>OR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urinary incontinence</td>
<td>32(33.3)</td>
<td>20(20)</td>
<td>1.67</td>
<td>1.02-2.70</td>
<td>0.035*</td>
<td>56.38</td>
<td>18.57-226.99</td>
<td>&lt;0.0001*</td>
</tr>
<tr>
<td>Incomplete bladder emptying</td>
<td>4(4.2)</td>
<td>10(10)</td>
<td>0.42</td>
<td>0.14-1.28</td>
<td>0.1</td>
<td>6.02</td>
<td>1.07-35.46</td>
<td>0.041*</td>
</tr>
<tr>
<td>straining</td>
<td>8(8.3)</td>
<td>11(11)</td>
<td>0.76</td>
<td>0.32-1.8</td>
<td>0.53</td>
<td>3.30</td>
<td>0.76-13.1</td>
<td>0.11</td>
</tr>
<tr>
<td>poor stream</td>
<td>4(4.2)</td>
<td>5(5.1)</td>
<td>0.82</td>
<td>0.23-2.95</td>
<td>0.76</td>
<td>5.88</td>
<td>0.51-152.89</td>
<td>0.16</td>
</tr>
<tr>
<td>Frequency (&gt;10 voids per day)</td>
<td>4(4.1)</td>
<td>9(9.1)</td>
<td>0.45</td>
<td>0.14-1.42</td>
<td>0.16</td>
<td>3.99</td>
<td>0.44-40.03</td>
<td>0.21</td>
</tr>
<tr>
<td>Nocturia</td>
<td>44(45.4%)</td>
<td>48(49%)</td>
<td>0.92</td>
<td>0.69-1.25</td>
<td>0.61</td>
<td>0.54</td>
<td>0.08-3.46</td>
<td>0.52</td>
</tr>
<tr>
<td>Double/triple voiding</td>
<td>15(15.79%)</td>
<td>12(12.1)</td>
<td>1.30</td>
<td>0.64-2.64</td>
<td>0.46</td>
<td>1.55</td>
<td>0.38-5.50</td>
<td>0.52</td>
</tr>
<tr>
<td>Dysuria</td>
<td>4(4.1)</td>
<td>3(3.1)</td>
<td>1.35</td>
<td>0.31-5.86</td>
<td>0.69</td>
<td>0.68</td>
<td>0.03-8.61</td>
<td>0.79</td>
</tr>
<tr>
<td>Urinary tract infection</td>
<td>9(9.3)</td>
<td>7(7.1)</td>
<td>1.31</td>
<td>0.51-3.38</td>
<td>0.57</td>
<td>0.85</td>
<td>0.14-4.01</td>
<td>0.84</td>
</tr>
</tbody>
</table>

*statistically significant P<0.05
7.2.3 Objective assessment of lower urinary tract function

Significantly more women who participated in clinical examinations compared with those who only answered the questionnaire had urinary incontinence symptoms (RR: 2.37, 95CI: 1.39-4.03 P: 0.0007) (Article 3).

In the pad-weighing test no difference was observed in urinary incontinence between SAH and TAH (Table 5).

A higher median functional bladder capacity was observed in the SAH group compared with the TAH group P= 0.0147 (Table 5). No other differences between SAH and TAH were found in the bladder diary (Table 5).

Urinary flow showed a higher median Qmax in the SAH group compared with the TAH group P= 0.042 (Table 5). We also found a trend towards a higher voided volume in the SAH group (Table 5). Figure 4 shows an increase in Qmax with increasing voided volume for SAH (blue) and TAH (red). The black line indicates the expected Qmax according to voided volume (the 50th percentile of the Liverpool nomogram44,47). The best logarithmic fits for our two surgical groups (red and blue lines) are slightly lower than the Liverpool nomogram but within normal range and all of them with very similar slopes. There was no difference in significant residual urine (Table 5).
Table 5 Pad-weighing, urinary flow, POP-Q and bladder diary according to type of hysterectomy. Intention-To-Treat.

<table>
<thead>
<tr>
<th></th>
<th>SAH</th>
<th>TAH</th>
<th>RR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pad-weighing test (SAH:53, TAH: 47), N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Objective UI</td>
<td>14 (26.4)</td>
<td>14 (29.8)</td>
<td>0.89</td>
<td>0.47-1.66</td>
<td>0.71</td>
</tr>
<tr>
<td>Subjective UI, Study questionnaire</td>
<td>21 (39.6)</td>
<td>16 (34.04)</td>
<td>1.16</td>
<td>0.69-1.95</td>
<td>0.56</td>
</tr>
<tr>
<td>Subjective and objective UI</td>
<td>9 (16.98)</td>
<td>7 (14.89)</td>
<td>1.14</td>
<td>0.46-2.82</td>
<td>0.78</td>
</tr>
<tr>
<td><strong>Urinary flow (SAH:51/TAH:44)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qmax median (interquartile range) (ml/sec)</td>
<td>30.3 (24.5-46.1)</td>
<td>27.1 (16.6-35.9)</td>
<td>0.89</td>
<td>0.47-1.66</td>
<td>0.71</td>
</tr>
<tr>
<td>Voided volume median (interquartile range) (ml)</td>
<td>415 (242.5-675.5)</td>
<td>307 (185.5-495.5)</td>
<td>0.89</td>
<td>0.47-1.66</td>
<td>0.71</td>
</tr>
<tr>
<td>Residual urine median (interquartile range) (ml)</td>
<td>0 (0-33.5)</td>
<td>0 (0-26)</td>
<td>1.14</td>
<td>0.46-2.82</td>
<td>0.78</td>
</tr>
<tr>
<td>Significant residual urine (&gt;100ml)* N (%)</td>
<td>5 (9.4%)</td>
<td>3 (6.7%)</td>
<td>1.14</td>
<td>0.46-2.82</td>
<td>0.78</td>
</tr>
<tr>
<td><strong>Pelvic organ prolapse N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Objective POP-Q</td>
<td>31 (59.6)</td>
<td>33 (70.2)</td>
<td>0.84</td>
<td>0.63-1.36</td>
<td>0.27</td>
</tr>
<tr>
<td>Objective (POP-Q) and subjective (symptoms), ( % of objective prolapse)</td>
<td>6 (19.4)</td>
<td>9 (27.3)</td>
<td>0.71</td>
<td>0.29-1.76</td>
<td>0.45</td>
</tr>
<tr>
<td><strong>Compartment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior</td>
<td>4 (7.55)</td>
<td>10 (21.3)</td>
<td>0.89</td>
<td>0.41-1.94</td>
<td>0.76</td>
</tr>
<tr>
<td>Posterior</td>
<td>16 (30.19)</td>
<td>13 (27.66)</td>
<td>1.09</td>
<td>0.59-2.02</td>
<td>0.78</td>
</tr>
<tr>
<td>Both</td>
<td>10 (18.9)</td>
<td>10 (21.3)</td>
<td>0.35</td>
<td>0.12-1.06</td>
<td>0.048</td>
</tr>
<tr>
<td>Apical (vaginal cuff/cervix)</td>
<td>1 (1.8)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bladder diary (SAH:34/TAH:34)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urinary incontinence* N (%)</td>
<td>9 (27.3)</td>
<td>5 (15.6)</td>
<td>1.7</td>
<td>0.66-4.65</td>
<td>0.25</td>
</tr>
<tr>
<td>Maximum voided volume (functional capacity) median (interquartile range) (ml)</td>
<td>531.7 (433.3-616.7)</td>
<td>443.3 (329.2-529.6)</td>
<td>0.73</td>
<td>0.41-1.94</td>
<td>0.76</td>
</tr>
<tr>
<td>24-hour fluid intake median (interquartile range) (ml)</td>
<td>1933 (1496-2507)</td>
<td>2085(1662-2653)</td>
<td>0.73</td>
<td>0.41-1.94</td>
<td>0.76</td>
</tr>
<tr>
<td>24-hour urine output median (interquartile range) (ml)</td>
<td>1910 (1577-2537)</td>
<td>1999 (1446-2518)</td>
<td>0.73</td>
<td>0.41-1.94</td>
<td>0.76</td>
</tr>
<tr>
<td>Nocturia* N (%)</td>
<td>10 (31.3)</td>
<td>14 (41.2)</td>
<td>0.75</td>
<td>0.4-1.46</td>
<td>0.4</td>
</tr>
<tr>
<td>Frequency* N (%)</td>
<td>8 (25)</td>
<td>13 (38.2)</td>
<td>0.65</td>
<td>0.31-1.35</td>
<td>0.25</td>
</tr>
<tr>
<td>Use of pads* N (%)</td>
<td>10 (31.3)</td>
<td>7 (21.9)</td>
<td>1.43</td>
<td>0.62-3.28</td>
<td>0.4</td>
</tr>
</tbody>
</table>

*χ² test  
bWilcoxon rank sum test  
cAccording to the follow-up questionnaire  
dDefined as waking up to void minimum once per night  
eDefined as voiding >8 times a day  
*Statistically significant P<0.05
Figure 3: Qmax according to voided volume for SAH and TAH

Black: Fitted according to the female Liverpool nomogram (50th percentile): 
\[ \ln(Q_{\text{max}}) = 0.511 + 0.505 \times \ln(\text{Voided volume}) \]
Red: Our best fit of TAH: 
\[ \ln(Q_{\text{max}}) = 0.3551668 + 0.5046498 \times \ln(\text{voided volume}) \]
Blue: Our best fit of SAH: 
\[ \ln(Q_{\text{max}}) = 0.6330179 + 0.4712442 \times \ln(\text{voided volume}) \]

7.3 Pelvic organ prolapse

According to the questionnaire 12 (12.9%) women had symptoms of pelvic organ prolapse after SAH; 11 (11.3%) reported this after TAH (P = 0.74) (Article 1).¹

There was a tendency towards more symptomatic women in the clinically examined population compared with those who only answered the questionnaire (RR: 2.19, 95%CI 0.94-5.08, P: 0.058) (Article 3).

POP-Q measurements showed that 31 (59.6%) in the SAH group and 33 (70.2%) in the TAH group had at least stage 2 pelvic organ prolapse; nevertheless, of the women with objective pelvic organ prolapse only 6/31 (19.4%) in the SAH group and 9/33 (27.3%) in the TAH group reported symptoms related to prolapse (Table 5).

¹ Article 1: Reference title or number.
A borderline difference between SAH and TAH regarding anterior pelvic organ prolapse was seen; less in the SAH group (N=4, 7.6%) compared with the SAH group (N=10, 21.3%) had anterior prolapse (RR: 0.35, 95% CI: 0.12-1.06, P=0.048) (Table 5).

7.4 Pain, constipation and hospital contacts

No differences between SAH and TAH were seen regarding pelvic pain or constipation (Table 6). Figure 4 shows hospital contacts/complications from time of hysterectomy to 14 years after. It shows that significantly more women had a bleeding related complication after TAH. After SAH 10 women had hospital contacts related to cervical problems, mainly continued bleeding. There is no overall difference in number of complications/hospital contacts between SAH and TAH.

7.5 Satisfaction with sex life

When comparing satisfaction with sex life after SAH and TAH for all responders we found no difference (Table 6). Those who stated they did not know if they were satisfied were excluded from the analysis (Article 1).1
Table 6: Secondary outcomes, Intention-To-Treat, Study questionnaire.

<table>
<thead>
<tr>
<th>Outcome (N= SAH/TAH)</th>
<th>Observed Data</th>
<th>Multiple Imputation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SAH</td>
<td>TAH</td>
</tr>
<tr>
<td>Constipation (N=97/100) (%)</td>
<td>14(14.4)</td>
<td>7(7)</td>
</tr>
<tr>
<td>Pelvic organ prolapse (N=93/97) (%)</td>
<td>12(12.9)</td>
<td>11(11.3)</td>
</tr>
<tr>
<td>Satisfied with sexual life (N=75/78) (%)</td>
<td>48(64)</td>
<td>53(67.9)</td>
</tr>
<tr>
<td>Pelvic pain (N=96/100) (%)</td>
<td>14(14.6)</td>
<td>10(10)</td>
</tr>
<tr>
<td>Vaginal bleeding (SAH only =97) (%)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Quality of life&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Component Score (PCS) mean [95% CI]</td>
<td>50.4 [48.5-52.4]</td>
<td>51.3 [49.4-53.2]</td>
</tr>
<tr>
<td>Mental Component Score (MCS) mean [95% CI]</td>
<td>54.8 [52.9-56.7]</td>
<td>53.2 [51.4-55.1]</td>
</tr>
<tr>
<td>Mean [95%CI]</td>
<td>SAH</td>
<td>50.05[48.5-51.6]</td>
</tr>
</tbody>
</table>

<sup>a</sup>Relative Risk  
<sup>b</sup>Statistically significant  
<sup>c</sup>Wilcoxon rank sum test
Figure 4: Hospital contacts from time of hysterectomy to 14 years

![Hospital contacts chart]

*P<0.05
7.6 Quality of life

Neither the physical (PCS) nor the mental (MCS) score of the SF-36 questionnaire differed between the SAH and TAH (Table 6). The means in our population were consistent with the means as defined by quality metric. When quality of life was assessed according to urinary incontinence we found that being urinary incontinent was associated with a lower mean physical quality of life score (48.0) than being continent (51.9), P=0.029 (wilcoxon’s rank sum test) (Article 2).

7.7 Cervical/vaginal screening and pathology

Data were obtained on 501 women (RCT N=317 (SAH: 157, TAH: 160)), (OS N=184 (SAH: 102, TAH 82)) (Figure 5). Seven from the SAH group and three from the TAH group had follow-up time of less than three years (Article 4).

The participants in the RCT and the OS did not differ in any characteristics other than mean follow-up time, which was longer for the RCT (Table 7). More women in the TAH group than in the SAH group were smokers at time of surgery.

After SAH 25 (9.7 %) of the women were never invited for screening. There was no difference in number of reminders sent to the women from the SAH group compared with the TAH group (Article 4).

From the SAH group 49 (18.9%) women actively withdrew from the screening program; 23 (8.9%) women gave hysterectomy as the reason for withdrawal. In the TAH group 191(78.9%) women withdrew from screening; 155 (64%) women withdrew because of hysterectomy (Article 4).

In the SAH group 159 (61.4%) women were adherent to the screening program and 78 (30.1%) in the SAH group had smears taken less frequently than every 5 years. In the TAH group 16 (6.6%) were
adherent and 67 (27.7%) had smears taken, but less frequently than every five years; in total 83 (34.3%) of the TAH group had minimum one vaginal vault smear taken after hysterectomy (Article 4). A total of 29 women had one or more abnormal smears. They were primarily found in the SAH group (N=28, 10.8%). Only one abnormal smear was found in the TAH group; it showed atypical cytology. Re-test after one year was normal. No cervical cancers were found. All abnormal tests were followed up by relevant tests (Article 4).
Figure 5: Flow of participants in the study of cervical cancer screening and cervical/vaginal pathology.
Table 7: Regarding Cervical Cancer Screening: Characteristics of participants in the randomized clinical trial (RCT) and the observational study (OS) according to hysterectomy method.

<table>
<thead>
<tr>
<th>Variable</th>
<th>RCT (N=317)</th>
<th></th>
<th>OS (N=184)</th>
<th></th>
<th>RCT vs OS P</th>
<th>SAH vs TAH P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SAH (N=157)</td>
<td>TAH (N=160)</td>
<td>SAH (N=102)</td>
<td>TAH (N=82)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age&lt;sup&gt;a&lt;/sup&gt; (years) mean (range)</td>
<td>62.2 (45.8-82.2)</td>
<td>61.6 (42.8-83.2)</td>
<td>63.0 (48.1-78.8)</td>
<td>62.3 (51.5-79.0)</td>
<td>0.15</td>
<td>0.11</td>
</tr>
<tr>
<td>Follow-up time&lt;sup&gt;b&lt;/sup&gt; (years) mean (range)</td>
<td>14.3 (0-17.8)</td>
<td>14.0 (0-17.7)</td>
<td>13.2 (0-16.8)</td>
<td>14.0 (0-16.3)</td>
<td>0.011*</td>
<td>0.26</td>
</tr>
<tr>
<td>Parity mean (range)</td>
<td>1.75 (0-5)</td>
<td>1.82 (0-4)</td>
<td>1.8 (0-6)</td>
<td>1.5 (0-3)</td>
<td>0.23</td>
<td>0.72</td>
</tr>
<tr>
<td>Smoking &gt; 5 cigarettes per day&lt;sup&gt;d&lt;/sup&gt; (%)</td>
<td>42 (26.8)</td>
<td>60 (37.5)</td>
<td>27 (27.8)</td>
<td>28 (35.9)</td>
<td>0.86</td>
<td>0.02*</td>
</tr>
<tr>
<td>Alcohol &gt; 14 units per week&lt;sup&gt;d&lt;/sup&gt; (%)</td>
<td>15 (9.6)</td>
<td>13 (8.1)</td>
<td>17 (17.5)</td>
<td>6 (7.7)</td>
<td>0.13</td>
<td>0.093</td>
</tr>
<tr>
<td>BMI preoperatively (kg/m&lt;sup&gt;2&lt;/sup&gt;) mean (range)</td>
<td>26 (17.8-42.6)</td>
<td>25.2 (17.1-47.3)</td>
<td>25.4 (17.2-38.3)</td>
<td>25.6 (19-35.9)</td>
<td>0.75</td>
<td>0.32</td>
</tr>
<tr>
<td>BMI 14 years (kg/m&lt;sup&gt;2&lt;/sup&gt;) mean (range)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>27.2 (17.7-42.6)</td>
<td>26.4 (16.7-50.1)</td>
<td>25.8 (17.2-37.2)</td>
<td>25.7 (18.4-43.3)</td>
<td>0.08</td>
<td>0.39</td>
</tr>
</tbody>
</table>

| Indication for hysterectomy                    |          |          |             |          |              |               |
|                                               | Fibroids | Abnormal uterine bleeding | Dysmenorrhoea | Pelvic pain | Endometriosis | Other |
|                                               | 93 (59.3) | 50 (31.9) | 6 (3.8) | 8 (5.1) | 0 | 0 |
|                                               | 92 (57.5) | 54 (33.8) | 6 (3.8) | 5 (3.1) | 0 | 3 (1.9) |
|                                               | 65 (63.7) | 29 (28.43) | 2 (1.96) | 4 (3.9) | 1 (0.98) | 1 (0.98) |
|                                               | 35 (42.7) | 34 (41.5) | 3 (3.7) | 6 (7.3) | 2 (2.4) | 2 (2.4) |
|                                               | 0.38 | 0.74 | 0.72 | 0.49 | 0.023* | 0.50 |
|                                               | 0.054 | 0.16 | 0.69 | 0.97 | 0.52 | 0.08 |

<sup>a</sup>Mean age of participants at time of follow-up (1. February 2014)

<sup>b</sup>Follow-up time is calculated from date of surgery until the first of the following: Death, left Denmark, turned 65 years of age (end of screening) or February 2014

<sup>c</sup>BMI at follow-up available for 189 women from the RCT (SAH: 98, TAH:92 ) and 118 from the OS (SAH:63, TAH: 53)

<sup>d</sup>At time of surgery

*Statistically significant
7.8 Meta-analyses of long-term follow-ups

Four trials of SAH versus TAH with more than 5 years follow-up, including our own trial were found and included in the meta-analyses (Table 8). In total the trials included 566 women (SAH: 286/TAH: 280) in questionnaire follow-up and 360 (SAH: 195, TAH: 165) in clinical examinations. The four trials with long term follow-up have participation rates of 65%, 27%, 82% and 64.3% respectively. The follow-up time was: 9.04 years, 9.1 years, 11.3 years and 14.1 years respectively and mean age at follow-up was 52.7, 51.3, 57 and 60.1 years. The data from our trial (Articles 1-3) and Persson et al were analyzed as intention-to-treat. This was not clearly stated in the trials by Greer et al and Thakar et al. The actual number included in each analysis is stated in the figure.

Urinary incontinence was addressed in all trials; however, overall urinary incontinence is only reported in our trial and the trial by Persson et al. The urinary incontinence subtypes stress and urgency incontinence were reported in all trials. The trials reported the number of women with stress or urgency urinary incontinence. Women classified as having mixed urinary incontinence in our trial were included in the meta-analysis of both stress and urgency urinary incontinence as they had symptoms of both. A meta-analysis of mixed urinary incontinence was not possible as the other trials did not state how many had both stress and urgency incontinence symptoms. The frequency of symptoms used to define a woman as being stress or urgency urinary incontinent varied between trials. Our trial included any answers of “in which situations do you experience urinary leakage” regardless of how often it occurred, Persson et al defined urinary incontinent as at least one leakage per month. Thakar et al used the categories 1) never, 2) occasionally, 3) weekly, 4) always. An answer of 2) or more was defined as urinary incontinent. Greer et al had the following categories: Never, 1-2 times per month, 1-2 times per week or daily. Any answer other than “never” is included as incontinent in the meta-analysis.

The analysis of stress urinary incontinence (Figure 6) showed more incontinence after SAH (OR: 1.61, 95%CI...
1.14-2.28). For urgency urinary incontinence (Figure 7) there was no difference between SAH and TAH.

The meta-analysis of the number of women with pelvic organ prolapse after SAH and TAH included only three trials: Greer et al\textsuperscript{50} did not include POP-Q examination and therefore is not included in this meta-analysis.

Pelvic organ prolapse according to POP-Q (minimum stage 2) did not differ between SAH and TAH (Figure 8). The heterogeneity between trials was low in all three meta-analysis. This was confirmed by the $\chi^2$ and I\textsuperscript{2} tests of heterogeneity (Figures 6-8).
Table 8: Randomized clinical trials of SAH and TAH with more than 5-year follow-up on clinical outcomes

<table>
<thead>
<tr>
<th>Author and title</th>
<th>N</th>
<th>Follow-up</th>
<th>Outcomes</th>
<th>Main findings</th>
<th>Strengths</th>
<th>weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thakar et al9</td>
<td>Questionnaire: 181</td>
<td>9.04 years</td>
<td>Quality of life (SF-36)</td>
<td>No difference between surgical groups.</td>
<td>Blinded Objective and subjective outcomes</td>
<td>Participation rate of 65% but less in the objective examinations. No handling of missing data</td>
</tr>
<tr>
<td></td>
<td>TAH: 90</td>
<td></td>
<td>Sexuality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SAH: 91</td>
<td></td>
<td>UI (stress and urgency separately)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Clinical examination:132</td>
<td></td>
<td>POP (POP-Q)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SAH:72</td>
<td></td>
<td>Sexuality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TAH:60</td>
<td></td>
<td>(questionnaire)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persson et al2</td>
<td>Questionnaire: 151</td>
<td>11.3 years</td>
<td>POP: POP-Q Pelvic floor symptoms (PFDI-20 questionnaire) UI + subtypes (interview)</td>
<td>No difference between surgical groups regarding pelvic floor symptoms or pop</td>
<td>High participation Objective and subjective outcomes</td>
<td>No handling of missing data</td>
</tr>
<tr>
<td></td>
<td>SAH: 80</td>
<td></td>
<td>TAH: 71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Clinical examination: 128</td>
<td></td>
<td>TAH: 58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SAH: 70</td>
<td></td>
<td>TAH: 58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greer et al8</td>
<td>Questionnaire</td>
<td>9.1 years</td>
<td>Pelvic symptoms, sexual function and quality of life (questionnaire)</td>
<td>No differences between groups</td>
<td>High loss-to-follow-up, only participation of one of the centres originally included in the trial. Insufficient power to reveal differences in UI No handling of missing data</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andersen et al (Article 1)1</td>
<td>197</td>
<td>14.1 years</td>
<td>Primary: UI (often or always), secondary: POP, pain, constipation, QoL, satisfaction with sex-life, hospital contacts and vaginal bleeding</td>
<td>More UI after SAH than after TAH, no other significant differences</td>
<td>Long follow-up Uses Multiple Imputation to account for missing data</td>
<td>Participation rate 64.8% Difference in results of observed data and MI data could indicate bias due to an unknown confounder Subjective outcomes only</td>
</tr>
<tr>
<td></td>
<td>SAH: 97</td>
<td></td>
<td>TAH: 100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andersen et al (Article 2):</td>
<td>197</td>
<td>14.1 years</td>
<td>Exploratory analyses of UI subtypes and LUTS as well as multivariate analyses of factors associated with UI and LUTS</td>
<td>Difference in UI between SAH and TAH is due to a difference in SUI BMI&gt;25 and properate UI are related to UI No difference in any other LUTS</td>
<td>Long follow-up Considers other factors that influence UI</td>
<td>Participation rate 64.8% Exploratory analyses Risk of mass significance because of many analyses No handling of missing data</td>
</tr>
<tr>
<td></td>
<td>SAH: 97</td>
<td></td>
<td>TAH: 100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andersen et al (Article 3):</td>
<td>100</td>
<td>14.02 years</td>
<td>Objective assessment of the following: POP: (POP-Q) UI (pad-weighing test) Voiding function (urinary flow and bladder diary) Bother related to the pelvic floor (PFDI-20 questionnaire)</td>
<td>No difference in objective UI and POP 60-70% had grade 2 POP – most without symptoms Higher Qmax and voided volume after SAH</td>
<td>Long follow-up Objective outcomes</td>
<td>Participation rate 33% No handling of missing data The participants have more UI problems than non-participants which may lead to more adverse outcomes than in the entire population</td>
</tr>
<tr>
<td></td>
<td>SAH: 53</td>
<td></td>
<td>TAH: 47</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 6: Meta-analysis of stress urinary incontinence in long-term follow-ups\(^a\) after SAH (experimental) versus TAH (control) abdominal hysterectomy

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Experimental SAH Events</th>
<th>Total</th>
<th>Control TAH Events</th>
<th>Total</th>
<th>Weight</th>
<th>Odds Ratio M-H, Fixed, 95% CI</th>
<th>Odds Ratio M-H, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andersen LL</td>
<td>60</td>
<td>97</td>
<td>44</td>
<td>100</td>
<td>32.7%</td>
<td>2.06 [1.17, 3.65]</td>
<td></td>
</tr>
<tr>
<td>Greer</td>
<td>9</td>
<td>18</td>
<td>5</td>
<td>19</td>
<td>4.8%</td>
<td>2.80 [0.71, 11.10]</td>
<td></td>
</tr>
<tr>
<td>Persson</td>
<td>27</td>
<td>70</td>
<td>17</td>
<td>58</td>
<td>22.6%</td>
<td>1.51 [0.72, 3.18]</td>
<td></td>
</tr>
<tr>
<td>Thakar</td>
<td>53</td>
<td>89</td>
<td>50</td>
<td>89</td>
<td>40.0%</td>
<td>1.15 [0.63, 2.08]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>274</td>
<td>266</td>
<td>100.0%</td>
<td>1.61 [1.14, 2.28]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total events 149/116
Heterogeneity: Chi\(^2\) = 2.62, df = 3 (P = 0.45); I\(^2\) = 0%
Test for overall effect: Z = 2.69 (P = 0.007)

\(^a\)Long-term Follow-up defined as more than 5 years of follow-up.

Figure 7: Meta-analysis of urgency urinary incontinence in long-term follow-ups\(^a\) after SAH (experimental) versus TAH (control) abdominal hysterectomy.

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Experimental SAH Events</th>
<th>Total</th>
<th>Control TAH Events</th>
<th>Total</th>
<th>Weight</th>
<th>Odds Ratio M-H, Fixed, 95% CI</th>
<th>Odds Ratio M-H, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andersen LL</td>
<td>31</td>
<td>97</td>
<td>35</td>
<td>100</td>
<td>43.9%</td>
<td>0.87 [0.48, 1.58]</td>
<td></td>
</tr>
<tr>
<td>Greer</td>
<td>8</td>
<td>18</td>
<td>6</td>
<td>19</td>
<td>6.1%</td>
<td>1.73 [0.45, 6.63]</td>
<td></td>
</tr>
<tr>
<td>Persson</td>
<td>10</td>
<td>70</td>
<td>7</td>
<td>58</td>
<td>12.3%</td>
<td>1.21 [0.43, 3.42]</td>
<td></td>
</tr>
<tr>
<td>Thakar</td>
<td>35</td>
<td>86</td>
<td>34</td>
<td>86</td>
<td>37.7%</td>
<td>1.05 [0.57, 1.93]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>271</td>
<td>263</td>
<td>100.0%</td>
<td>1.03 [0.71, 1.51]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total events 84/82
Heterogeneity: Chi\(^2\) = 0.98, df = 3 (P = 0.81); I\(^2\) = 0%
Test for overall effect: Z = 0.17 (P = 0.86)

\(^a\)Long-term Follow-up defined as more than 5 years of follow-up.
Figure 8: Meta-analysis of grade 2 prolapse\(^b\) in long-term follow-ups\(^a\) after SAH (experimental) versus TAH (control).

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Experimental SAH</th>
<th>Control TAH</th>
<th>Odds Ratio M.H, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andersen LL clinical exams</td>
<td>31</td>
<td>53</td>
<td>0.60 [0.28, 1.37]</td>
</tr>
<tr>
<td>Persson</td>
<td>27</td>
<td>70</td>
<td>1.03 [0.50, 2.10]</td>
</tr>
<tr>
<td>Thakar</td>
<td>4</td>
<td>65</td>
<td>0.61 [0.15, 2.20]</td>
</tr>
</tbody>
</table>

Total (95% CI) 188 167 100.0% 0.78 [0.47, 1.29]
Total events 62 61
Heterogeneity: CHI\(^2\) = 1.09, df = 2 (P = 0.50); P = 0%
Test for overall effect: Z = 0.97 (P = 0.33)

\(^a\)Long-term Follow-up defined as more than 5 years of follow-up.

\(^b\)Measured by POP-
8. Discussion

8.1 Overall findings

In this long-term follow-up, we found that more women reported urinary incontinence after SAH than after TAH; the difference was mainly due to a difference in the frequency of stress urinary incontinence. We observed no difference in the secondary outcomes: Pain, constipation, satisfaction with sex life, pelvic organ prolapse, and quality of life. There was no difference in number of hospital contacts between surgical groups.

Regarding the exploratory analyses no difference was found between surgical groups regarding any other LUTS. Preoperative urinary incontinence, BMI and local oestrogen treatment were associated with urinary incontinence. BMI was primarily associated with urgency symptoms and mixed urinary incontinence. Furthermore, we found no difference in objectively assessed urinary incontinence or pelvic organ prolapse. A higher Q-max and functional bladder capacity was seen after SAH. Adherence to cervical cancer screening was 61.4% after SAH and approximately 90% had at least one cervical smear done after SAH. Nevertheless almost 10% were not invited for cervical cancer screening after SAH. Abnormal cervical smears were found in 10.8% of the women after SAH. In the TAH group 34.3% had at least one vaginal vault smear done in the follow-up period. We found no cervical cancers.

8.2 Strengths

The main strengths of the studies in this thesis are: They are based on a randomized clinical trial (RCT); it is the largest RCT comparing SAH with TAH and has the longest follow-up. Furthermore
we have included subjective and objective clinical outcomes as well as registry data allowing a full comparison of the two surgical methods.

8.2.1 Randomized clinical trial

The randomized clinical trial design is a strength because it ensures that the SAH and TAH groups were in fact comparable at the beginning of the trial regarding known and unknown confounders. Differences found between groups can be interpreted as associated with the previous hysterectomy. If they were not randomized one category of patients could be more likely to be offered one kind of treatment and this might affect the results of the treatment. The RCT design reduces this risk of selection bias by random sequence generation and allocation concealment.

8.2.2 Long-term follow-up

The long follow-up is a strength because it allows us to compare consequences of hysterectomy that are not apparent until later in life, the women in the trial are now post-menopausal. The outcomes of interest are known to increase after menopause meaning that the possibility of detecting a difference between SAH and TAH, if there truly is a difference, will increase.

8.2.3 Sample size

The larger population is a strength because it increases the power of the trial meaning that it increases the likelihood that we reveal a true difference between SAH and TAH (i.e. reject the null-hypothesis). It also means that we can more readily accept the null-hypothesis in outcomes where we find no difference. A smaller sample size would be more likely to show non-significant results even if a true difference exists (Type 2 error).
8.2.4 Cervical screening/pathology

The strength of the registry based study of cervical cancer screening and pathology (Article 4) is that we have information on all the women included in the RCT and OS for a long period of time. Therefore, we are able to assess the occurrence of dysplasia/cancer of the cervix as well as the effectiveness of the cervical cancer screening program in Denmark in relation to women who have undergone hysterectomy in these cohorts.

8.3 Limitations

Some limitations should be considered when interpreting the results of this thesis: lack of blinding of participants and surgeons, loss to follow-up, risk of mass significance and sample size in regards to cervical pathology. Lastly, we have not validated the data from patobank against medical charts or other national registries, so the completeness of the registry could be questioned.

8.3.1 Blinding

A way of reducing bias in RCTs is blinding of participants and personnel, and blinding of outcome assessment. Not blinding the women in this trial was a decision based on ethical considerations. The women in the SAH group should know to continue cervical cancer screening and it would be unethical to blind them leaving them at higher risk of developing cervical cancer. Furthermore, some women experience vaginal bleeding or spotting after subtotal hysterectomy and would be unblinded by this. Self-examination or a pelvic exam conducted by a physician unaware of the trial could also easily unblind the participant.
8.3.2 Missing data

Another limitation of this thesis is the large drop-out leading to missing data. Long-term follow-ups whether conducted as RCTs or non-randomized cohort studies are prone to missing data. Missing data can arise because of loss to follow-up or because of incomplete follow-up i.e. questionnaires only partly completed. Missing data increases with duration of follow-up. Reasons for this can be that participants cannot be located, have lost interest and do not wish to participate/only participate in the parts they find interesting, or have died. Brubaker et al found that missing data were higher among younger participants and regarding data registered by participants when they were away from the research team — such as postal questionnaires.

The loss to follow-up might compromise the internal validity of the trial if an imbalance was seen in responders between intervention groups. However, the responders in the two surgical groups in this thesis are comparable according to baseline characteristics. Non-responders might differ in a systematic way from responders meaning that our sample at follow-up is less representative of the population of women undergoing hysterectomy for benign indications than the original trial sample. This could compromise the external validity of the trial. Besides smoking and alcohol consumption at time of surgery, our baseline data do not demonstrate systematic differences. However, factors unknown to us could also result in bias.

It seems that women who participated in clinical examinations differed from women who only filled-in the questionnaire; more had urinary incontinence among those who participated in clinical examinations. It is likely that the existence of pelvic floor and/or urinary tract symptoms were a motivational factor for participation for the women.

Explanations of our low response could be loss of interest due to the long follow-up, older age, or medical conditions making it overwhelming to be asked to participate. This is reflected in our high
return of the questionnaire: We were able to contact the women; however, some actively chose not to fill-in the questionnaire. Another possible reason is that according to the ethics committee, we were allowed to contact participants by letter only. Contact by telephone was allowed in earlier follow-ups and in a comparable trial. It might also have discouraged participants that we included too much in the follow-up (questionnaire, clinical examination and bladder diary). The fact that the questionnaire included questions of an intimate nature might also have discouraged participation.

Missing data can be categorised as: Completely at random (MCAR), missing at random (MAR), or missing not at random” (MNAR). MCAR means that data are lost by coincidence (questionnaire lost in the mail, machine for analyzing blood samples broken one day) this will not introduce bias; however, it will affect the power of the trial because of decreased sample size. MAR means that missing data is related to other observed variables, but not to unobserved variables nor to the missing variables themselves. An example of this is if more participants from one age group than from others drop out. In this case missingness could introduce bias, but we can account for it because the variable related to missingness is known to us and can be included in a statistical model. MNAR means that the missing value is related to the value itself or related to another variable that is not known or registered in the trial. An example of MNAR is smoking cessation studies where it is likely that non-response is related to smoking relapse. This constitutes a larger problem as this can introduce bias and there is no valid statistical method to account for this.

Several methods have been suggested to handle missing data: In complete case analysis all participants with missing data are excluded – simply ignoring the missing data. In carry forward analyses the missing data is replaced with the last known value of the outcome. Single imputation replaces the missing value with the mean of all the known values. These methods have all been used but have several limitations; standard errors become too small, the risk of attrition bias remains high.
and for complete case analysis the statistical power decreases because of a smaller sample size.\textsuperscript{55}

Multiple Imputation (MI) is a statistical method where a number of complete datasets are constructed based on the information available regarding data from the variable in question as well as other related variables. Multiple Imputation is based on the assumption that missingness is MAR.\textsuperscript{56}

In our analyses of the primary and secondary outcomes \textit{(Article 1)}\textsuperscript{1} we used MI to handle missing data. These analyses showed no significant difference regarding urinary incontinence or any other outcomes. This might reflect a true non-significant difference and our finding of more urinary incontinence after SAH may be caused by attrition bias. It could also reflect the uncertainty incorporated in MI (within imputation variance and between imputation variance)\textsuperscript{54} which increases with the percentage of missing data — in this case leading to non-significant results. A third explanation of the difference between our observed and MI results could be an unknown confounding factor related to missingness as well as urinary incontinence meaning that our ‘missing at random’ (MAR) assumption is incorrect and the data is truly ‘missing not at random’ (MNAR). In this case the MI analysis is not valid.\textsuperscript{54} Our observed findings corresponded well with prior follow-ups of this trial with higher response proportions.\textsuperscript{21,22} The relative risk found in MI for urinary incontinence is equal to the relative risk in the 5-year follow-up,\textsuperscript{22} however, in the 5-year analysis the P-value was 0.052 — much closer to a statistically significant result. This could reflect the increased uncertainty because of the higher percentage of missing data in the current follow-up as explained above. It is also important to keep in mind that multiple imputation datasets are constructed datasets based on qualified guesses of the true values of the missing data. The quality of the guess depends on the factors included in the model and how accurately they can predict the missing values.
In the clinical examinations (Article 3) our loss to follow-up was about 67%. The results from the clinical examinations can be interpreted as subgroup analyses of the women who most likely have the most pelvic floor and urinary tract symptoms as indicated by their answers to the study questionnaire. Major conclusions should not be drawn from these results because of the low participation; however, they can raise questions and encourage further research.

8.3.3 Mass significance

Another limitation is the risk of mass significance due to the high number of outcomes. We have defined a significance level of 5%. This means that we could expect that one in twenty comparisons would be statistically significant just by chance (Type 1 error). We do not believe that our finding of more urinary incontinence after SAH is due to mass significance: This finding is consistent with findings in earlier follow-ups (Figure 2); it is unlikely that a chance finding due to mass significance would be reproducible at all follow-up times. Further, urinary incontinence was the primary outcome which was pre-specified at the beginning of the trial. The sample size was based on the ability to detect a difference in urinary incontinence at the 5% significance level.

8.3.4 Sample size

A limitation of the study regarding cervical cancer screening and pathology (Article 4) is the length of the follow-up and the population size. Cervical stump cancer is rare which means a large study population is needed to gain valid information about the incidence of cervical cancer after subtotal hysterectomy. Our population might be too small and our 14-year follow-up might be too short to evaluate this risk.
8.4 Interpretation

8.4.1 Urinary incontinence

The difference in the frequency of urinary incontinence we found is consistent with findings at one and five-year follow-ups of the trial.\(^{21,22}\) The percentage of women with subjective urinary incontinence has increased in both surgical groups over the years; however, more so in the SAH group. This result seems counter intuitive as most would expect that the less invasive and nerve sparing procedure i.e. SAH would result in the best long term outcome. The results could reflect a difference in treatment seeking behaviour rather than a difference in occurrence of urinary incontinence. The women in the TAH group could be more likely to be treated for their incontinence than the women in the SAH group resulting in more current urinary incontinence in the SAH group. However, the analyses including “urinary incontinence at any follow-up” showed the same difference (Article 1).\(^1\)

The difference in subjective urinary incontinence is not consistent with findings in other long-term follow-ups.\(^{49-51}\) However, consistent with our finding the meta-analysis we conducted of stress urinary incontinence showed a benefit of TAH compared with SAH. Also consistent with the findings in our trial, the meta-analysis of urgency urinary incontinence showed no difference between surgical groups. Persson et al\(^{51}\) concludes that most outcomes are in favour of total hysterectomy although they do not reach statistical significance. They suggest that longer follow-up and a larger population might lead to significant differences. Our follow-up is longer and we have more participants; this might explain why we, contrary to the other trials find a significant difference in urinary incontinence. Thakar et al\(^{49}\) found no difference between SAH and TAH regarding urinary incontinence. The prevalence of urinary incontinence differs between ethnic groups; stress urinary incontinence is less common in women of
African descent. This could explain the difference between our and Thakar et al's results; their population was mixed and one third of the population was of African descent while ours was homogenous Caucasian. The study by Greer et al included only 37 women from one of the centers originally participating in the trial. They found no difference between surgical groups; the sample size is clearly too small to reveal a difference in urinary incontinence and there is a high risk of attrition bias which was not addressed in the paper.

The difference in stress urinary incontinence between SAH and TAH could be explained by a difference in suspension of the vagina at surgery. A method of suspension to be used was not specified for this trial; however, it was common to perform suspension of the vagina (in most cases from the vaginal top to the cardinal and/or the sacrouterine ligaments) during TAH but not during SAH. The suspension performed during TAH has been suggested to yield more support and decrease bladder neck mobility compared with the native tissue of SAH.

We could not reproduce the difference in urinary incontinence according to the questionnaire in the pad-weighing test or according to the bladder diary. This could be because of an insufficient sample size to reveal a difference. It could also be because the subgroup that participated in clinical examinations was not representative of the entire study population; they had more subjective urinary incontinence than those who only answered the questionnaire.

High BMI was significantly associated with urinary incontinence. One could argue that women with stress urinary incontinence tend to avoid physical activity that might lead to leakages and become overweight as a result of lack of exercise. Conversely, the higher intra-abdominal pressure associated with overweight might be strenuous to the pelvic floor and could be the reason for incontinence.

Nonetheless, we found that urgency symptoms and mixed urinary incontinence were associated with...
higher BMI, whereas stress urinary incontinence was not. Khullar et al.\textsuperscript{61} found an association between higher BMI and all urinary incontinence subtypes. In keeping with our results, Townsend et al.\textsuperscript{60} found that increasing BMI was associated with higher incidence of urgency and mixed, but not stress incontinence. Nevertheless, they found that increasing waist circumference was associated with higher incidence of stress but not urgency and mixed incontinence. Consequently, if we had measured waist circumference rather than BMI, the result might have been different.

It has been shown\textsuperscript{62} that weight loss significantly improves urinary incontinence, especially stress incontinence, in overweight women. Moreover, a recent study from New Zealand\textsuperscript{63} indicates that the metabolic changes following bariatric surgery, such as normalization of insulin resistance, are more important than actual weight loss when it comes to improvement of LUTS, especially urinary incontinence. The improvement is seen only 6 weeks after surgery, before the major weight loss. In keeping with this, a recent systematic review\textsuperscript{64} concluded that current evidence suggests a link between overactive bladder and metabolic syndrome. This shows that the link between obesity and urinary incontinence and LUTS remains unclear.

8.4.2 Urinary Flow

We found a higher Qmax and voided volume in the SAH group compared with the TAH group. The relationship between Qmax and voided volume is well known\textsuperscript{65} and described in the Liverpool nomogram.\textsuperscript{44} The women were asked to present with a comfortably full bladder and our results indicate a larger bladder capacity after SAH compared with TAH. This is consistent with the bladder diary where the SAH group had a higher functional capacity compared with the TAH group. This could explain the higher proportion of women in the SAH group with urinary incontinence symptoms.\textsuperscript{1} If the women in the SAH group tend to have a bladder with higher volume, they might be more likely to
leak. To confirm a larger bladder capacity in the SAH group full urodynamics would have to be conducted. We chose not to do this as we anticipated lower participation if we had included invasive urodynamics.

8.4.3 Pelvic organ prolapse

We found no difference between SAH and TAH regarding pelvic organ prolapse symptoms according to the questionnaire or objective pelvic organ prolapse according to POP-Q except more objective anterior prolapses in the TAH group; however, the numbers are so small and the result is of borderline significance so such a result should be interpreted cautiously. The meta-analysis of three RCTs with long-term follow-up including POP-Q measurement showed no significant difference in pelvic organ prolapse between SAH and TAH.

A smaller proportion stated they had pelvic organ prolapse in the study questionnaire than what we found in the objective examinations. Persson et al. did not state how many had subjective pelvic organ prolapse and Thakar et al. found that a higher percentage had feelings of pelvic organ prolapse than actual stage 2 pelvic organ prolapse.

Several reasons can be identified for the discrepancy between the results from the study questionnaire and the clinical examinations: Many of the stage 2 pelvic organ prolapses we found in our clinical examinations were asymptomatic and of no significance to the woman. Furthermore, a Swedish study showed that one in five women with symptomatic pelvic organ prolapse were not aware that the symptoms were caused by the prolapse. This means that even though they have pelvic floor symptoms they are not aware that they have a prolapse. A recent study showed that the only symptom of pelvic organ prolapse that correlates well with POP-Q measurement is the feeling of a bulge in the vagina.
8.4.4 Cervical cancer screening

One in ten women was never invited to participate in the national screening program for cervical cancer after SAH. These women still have a cervix and therefore an indication for cervical cancer screening.

In the SAH group 8.5% of the women did not have a smear taken in the follow-up period. They were at increased risk of a late diagnosis of cervical pre-cancer or cancer. In Denmark the overall cervical screening compliance is about 66.6%. We found that 61.7% were adherent to the screening program. So even though some women are not invited after SAH the participation rate is comparable to that of the overall population.

More than 90% of the women in the SAH group had at least one smear taken in the follow-up period. Nonetheless, almost 20% withdrew from the screening program, half of whom gave hysterectomy as the reason. This suggests that the importance of continued screening after SAH is not clear to all the women who undergo subtotal hysterectomy. All of the women in the SAH group were informed at entry into the trial to continue screening after hysterectomy. Whether the discontinuation is due to wrong advice from a doctor later on or a misunderstanding by the participant is unknown.

One third of the TAH group had at least one smear. Cervical screening programs effectively reduce the number of invasive cervical cancers; however, it is costly, can induce anxiety, and should be used appropriately. In the US, regular vaginal smears after total hysterectomy have not been recommended since 2003; however, a study from 2010 showed that about 60% of these women still had vaginal smears done. A British NHS national guideline recommends the following program regarding vaginal vault smears after total hysterectomy: If the woman has had regular cervical smears for 10 years prior to hysterectomy and no abnormal smears there is no indication for further screening. If the
woman has had regular smears without abnormal findings for less than 10 years prior to hysterectomy. It is suggested that one vaginal vault smear should be taken 6 months after total hysterectomy. If any smears taken prior to hysterectomy were abnormal further smears are recommended.

An inclusion criteria for participation in our RCT and OS was no previous abnormal pap-smear.\textsuperscript{21} Further, all women in our trial had a normal pap-smear at entry into the trial. According to the NHS criteria there should be no reason for continued follow-up of our participants. Because of the lack of a guideline in Denmark, the women and the general practitioners might be uncertain whether the smear is indicated or not—especially when the women continue to receive invitations as was the case in our study. A study from the UK\textsuperscript{72} showed that greater knowledge among general practitioners and practice nurses was associated with fewer vaginal vault smears taken. This might also be the case in Denmark. The lack of a guideline might be leading to excessive testing under the “better safe than sorry” mantra.

We found no cervical cancers in our follow-up. Cervical stump cancers are seen in 1-3\% of women who have undergone subtotal hysterectomy\textsuperscript{30} which means that if our follow-up time is long enough we would expect to find 2.5-7.7 cases in our material. Nevertheless, our participants might be less likely to develop cervical stump cancer because they were excluded from participation if they had any prior abnormal smears. Furthermore, Hellström et al\textsuperscript{30} found that the time interval from subtotal hysterectomy until diagnosis of cervical stump cancer was 17.6 years on average.
9. Conclusion

The three studies included in this thesis all compare long-term consequences of subtotal and total abdominal hysterectomy in a Danish, Caucasian population. We found:

- No results favouring the SAH procedure regarding long-term outcomes.
- SAH was inferior to TAH regarding urinary incontinence – especially stress urinary incontinence.
- Cervical cancer screening is performed inadequately after hysterectomy.
- One in ten experience abnormal smear results after SAH leading to further tests and interventions.

10. Recommendations

Short and long term benefits should be considered when recommending one method of hysterectomy over another. The shorter surgical time, less blood loss and maybe fewer short term complications associated with subtotal hysterectomy must be weighed against the long term consequences of an increased risk of urinary incontinence and the risk of cervical pathology. Based on this thesis we recommend that total hysterectomy is the first choice for benign diseases of the uterus.
11. Future aspects

Surgical methods for hysterectomy are changing faster than we can produce clinical trials to compare outcomes. However, these trials are still an important tool for evaluating and comparing surgical methods resulting in evidence based recommendations.

In spite of the difficulties associated with long-term follow-up, long-term outcomes should be assessed as these may develop years after surgery — especially after menopause. Laparoscopic hysterectomy is used increasingly among women who have to have their uterus removed for benign reasons. This study concludes regarding abdominal hysterectomies and not laparoscopic hysterectomy. Future research should focus on the minimally invasive surgical methods. Subtotal and total laparoscopic/robotic hysterectomy should be compared in RCTs.

The use of morcellation, which makes subtotal laparoscopic hysterectomy possible, has been associated with harmful consequences and is no longer recommended by the FDA\textsuperscript{6}. However, the advocates for using morcellation are testing new methods of contained morcellation\textsuperscript{73-75} which might prove to be safe.

The vaginal hysterectomy approach is also considered minimally invasive. This method is a fast and cheap alternative to the modern laparoscopic methods and more long-term RCTs comparing vaginal and laparoscopic hysterectomy are indicated. According to the Danish national guideline the vaginal hysterectomy should be considered when possible.\textsuperscript{12}

The results in our study of a difference in urinary incontinence could be due to a difference in vaginal/cervical suspension during subtotal and total hysterectomy. Research looking further into which methods of suspension are being used as well as randomized trials comparing methods of suspension at hysterectomy is warranted in order to shed more light on this.
The link between obesity and UI and other LUTS is still unclear and future research regarding the association between weight, metabolic disorders and LUTS is warranted.

We found that the cervical screening program is less than optimal in handling hysterectomized women. A nationwide study relying on national registries to fully evaluate the effectiveness of cervical screening in post hysterectomy women could increase our knowledge on this matter. A national guideline stating indications for post-hysterectomy smears after subtotal as well as total hysterectomy would probably decrease the number of unnecessary smears after total hysterectomy and increase adherence to screening after subtotal hysterectomy.
12. References


(10) DHHD. Danish Hysterectomy and Hysteroscopy database yearly report 2013. 2014.


(28) The Danish National Health authorities. HPV-Vaccination a part of the Danish Children's vaccination program. 2014.


13. Appendices

13.1 Questionnaire SAH

An English version of the study questionnaire for SAH can be seen here:
https://drive.google.com/file/d/0B5k2eI7IcEBDM20tUWQ0bmZWbEU/view?usp=sharing

13.2 Questionnaire TAH

An English version of the study questionnaire for TAH can be seen here:
https://drive.google.com/file/d/0B5k2eI7IcEBDQmNzWFRrT1R0bk0/view?usp=sharing

13.3 Questionnaire PFDI-20

The Danish version of the PFDI-20 questionnaire can be seen here:
https://drive.google.com/file/d/0B5k2eI7IcEBDcXJ3SEJCDcY4cTg/view?usp=sharing

13.4 Bladder Diary

The Danish version of the bladder diary can be seen here:
https://drive.google.com/file/d/0B5k2eI7IcEBDUnc1WkZyRXZycG8/view?usp=sharing

14. Articles