Best practices in minimally invasive gynecology: making sense of the evidence

Sandberg, E.M.

Chapter 13

General discussion

The implementation of minimally invasive surgery (MIS) in general and of laparoscopic hysterectomy (LH) in particular, has been accelerated in the field of gynecology over the past 30 years.[1-3] As this new surgical technique was introduced so rapidly, surgeons may have developed their own authority based management, which most probably resulted in medical practice variations amongst hospitals and/or surgeons. In a time period of increasing transparency and strive for standardization, this thesis aimed to formulate best practices for clinical topics of LH and laparoscopic myomectomy (LM) that were associated with limited scientific support.

To start, a guideline for LH was developed to standardize this procedure. Chapter 2 summarizes the clinical recommendations of the guideline. A considerable finding was the fact that with the use of the GRADE method the quality of available evidence was frequently graded as *low* or *very low*. Of the more than 60 outcomes studied in the guideline, it appeared that for 38 of them (63%), the quality of the evidence was very low. Inherently, this opens the debate regarding the added value and relevance of the formulated recommendations.

Implication of a guideline associated with low quality of evidence

Firstly, to understand how the evidence was graded, background information on the GRADE method is essential. The GRADE approach was developed at the beginning of the 21st century by the group that had also introduced the term *evidence-based medicine* twenty years earlier.[4] This GRADE tool is meant to systematically evaluate the quality of evidence. It differes from previous methods as it assesses the strength of the evidence for defined outcomes, rather than for individual studies.[4] For each outcome, the GRADE method suggests to systematically grade five domains, which then generates a final level of evidence (high, moderate, low and very low quality of evidence). Although this method is internationally one of the most recognized tools in medicine, it is important to realize that it was originally developed by physicians of internal medicine. This is particularly relevant to consider when looking at the definitions set by the GRADE working group regarding the level of evidence. Indeed, high quality of evidence is defined as 'it is unlikely that further research will have an effect on the outcome'. For research with medication for example, this definition seems plausible, as for the same group of people further research will most probably not provide new insights. Yet, for research with surgical techniques, this definition seems less appropriate since techniques, technologies and the experience of surgeons keep evolving. For instance, one of the first randomized controlled trials (RCTs) comparing LH to abdominal hysterectomy did not demonstrate relevant significant advantages in favor

of LH.[5] In contrast, LH was associated with a prolonged operative time and an increased risk of complications, particularly to the urinary tract.[5] Yet with the further development of the laparoscopic technique and the increasing experience of the surgeons, it became clear that the laparoscopic approach was superior to the abdominal one.[6] From a GRADE point of view, many outcomes in this first trial would have most probably been classified as *high quality*, signifying that it is unlikely that further research will have an effect on the outcome. Thus, to avoid negatively interfering with the further development of a surgical technique, it is essential to interpret the results of the GRADE method in the right context.

Another limitation of the GRADE approach is that it is primarily designed for RCTs. When non-RCTs are included in the framework, the level of evidence is immediately downgraded, leading to an overall *low* or *very-low* quality of evidence. Since the introduction of evidence-based medicine, RCTs have been considered as the highest and most reliable study design. Although RCTs do indeed have advantages (e.g. the elimination of confounders), RCTs are for certain outcomes not always appropriate. For example, for rare complications, RCTs are often underpowered and large cohort analyses are much more relevant, provided that potential bias is taken into account. In **chapter 3** we observed this as well: RCTs alone did not demonstrate a difference between LH and VH for the risk of vaginal cuff dehiscence, that was only found in the overall analysis. In general, in our opinion, doctors but also clinical researchers and guideline-developers should more structurally incorporate the outcomes from large databases and should for this purpose take advantages of the increasing possibilities of information technology (IT) and big data. In the same line, the GRADE method, but also the Cochrane group, should consider cohort analyses from (national) large well-designed databases as a valid research tool, rather than focusing on RCTs only. Therefore, an adapted GRADE tool should be developed for the evaluation of surgical innovation.

The above points of criticism towards the GRADE method have been raised previously.[7] The GRADE working group has replied to this criticism by stating that their proposed method should be considered as a framework to get systematically insight into the quality of the available evidence, and that the process of translating this evidence into recommendations should be an open discussion. The GRADE working group stated that 'different expert panels may come up with somewhat different strength of recommendations based on the same body of evidence given variation in their clinical and socioeconomic settings and population of interest, and valued judgments'.[8] The level of evidence is thus not the same as the strength of the recommendation.[8] This idea is underlined in this thesis when looking firstly at the policy regarding the standard use of a cystoscopy after LH. The guideline of the American Association of Gynecologic Laparoscopists (AAGL)[9] recommends

considering routine cystoscopy after every LH while in the Netherlands, based on the same literature, we advise selective use (**chapter 2** and **chapter 5**). A second example is observed in **chapter 4**, where it was demonstrated that, although the LESS technique for hysterectomy is feasible, safe and equally effective compared to the conventional laparoscopic approach, it has no clinically relevant advantages. Yet, in certain parts of the world, mainly in Asia, the laproendoscopic single site surgery (LESS) technique has been widely implemented. In certain Asian hospitals, up to 80% of the laparoscopic surgery is done in this fashion.[10] These examples also demonstrate that evidence-based medicine is not as objective as we always claim and hope it is.

Despite the understanding of the origin of the *low* quality of evidence and the interpretation of these results, concerns were raised during the development of the guideline of LH (**chapter 2**) regarding the medico-legal impact of such a document. Although this guideline might indeed be used for liability issues, it is in our opinion a strength to have an 'official' document issued from our national medical society where the minimal requirements for LH are summarized. Furthermore, from medico-legal point of view, it can be assumed that if no Dutch documents are available, experts in court will most probably base their verdicts on other (European) guidelines or available literature. In Germany, France and Denmark similar guidelines have already been developed, again based on the same literature.[11;12;13] Yet, as discussed previously, the strength of the recommendations and the considered topics may vary according to the settings. As a result, we believe that it is a must to have a guideline tailored to the Dutch practice.

Besides the medico-legal aspects, a national guideline for LH is of additional value for the professionals (e.g. the gynecologists) as it provides an overview of the best practices for the procedure at issue. With the additional research performed thereafter in this thesis, most actual topics of LH have been covered and standardized. However, the job is not finished: as the LH technique keeps evolving, new clinical challenges will most certainly be faced. Surgical innovation is a dynamic process and therefore it is essential, for the surgical field in general, to keep evaluating the outcomes of surgical procedures.

Implementation of a new surgical technique

As already mentioned in the introduction, implementing surgical innovations is one of the most complex dilemmas in medicine. In **chapter 10**, we studied the newly-introduced uterine-sparing techniques for fibroid treatment and demonstrated that, compared to the conventional approaches such as myomectomy and embolization, the re-intervention risk after High Intensity Focused Ultrasound (HIFU) was associated with the least promising

results. Also, for most recent introduced uterine-sparing techniques such as Radio-frequency Ablation (RFA), we observed that long-term data were lacking. It is important to realize that most techniques described in **chapter 10** were FDA approved and were, despite the lack of evidence, not limited to clinical trials anymore. As a result, systematic data collection remains essential and long-term data are urgently needed.

In the same line, the power morcellator is one of the instruments that was rapidly introduced and of which an important limitation was overlooked, partly because of lack of long-term surveillance. It seems that for this specific example, history is repeating itself. Indeed, after the press release of the Food and Drug Administration (FDA) in 2014,[14] gynecologists and medical device industries have sought to develop techniques to reduce the risk of potential spread while conserving the benefits associated with the laparoscopic approach. Contained tissue extraction has been suggested as solution and in many clinics this technique with a bag has been rapidly adopted. However, these bags are being introduced off-label and again without proper systematic evaluation prior to implementation. In addition, in the Netherlands at least, no national registration system is currently in place which will allow for data analyses over time. Yet, it is absolutely necessary to collect (national) data as it is questionable if this contained tissue extraction technique is for example for myomectomy safe from oncological point of view, as we discussed in **chapter 8**. Even during abdominal myomectomy without morcellation, micro-spillage of tissue was observed in the abdomen. Although the clinical relevance of tissue dissemination at this level is unclear, it cannot be concluded that it is harmless.

Ideally, the outcome measures of all (laparoscopic) surgical procedures in the Netherlands should be registered in a similar way as done in the Scandinavian countries or by the Dutch Institute for Clinical Auditing (DICA). An argument in favor of such large registration system was underlined in **chapter 9**, where we observed that for laparoscopic myomectomy (LM) of more than 500 grams, conversion risk significantly increased. These conclusions were based on data from a large center in the United States. Looking at the number of myomectomies performed yearly in the Netherlands, such conclusions could have only been drawn by collaborating between hospitals. A national registration system will give us more insight into our general performance and will allow us to make the necessary improvements. The DICA has already proven that their yearly audits significantly improve quality of care.[15] For the field of benign gynecology in the Netherlands, the current system in place does not result in structural data collection, in the first place as data collection is not done automatically and is not mandatory. A first step was made in April 2017 though, when the Dutch government passed a law obliging all medical implants to be registered in a national system.[16] Yet, an additional crucial step for a successful registration is in our

opinion, a simple and user-friendly system. We therefore plea for a high-tech registration system supported by professional ICT resources, even though this is a financial investment.

Evidence-based medicine: measuring outcome

Besides appropriate data collection, adequate interpretation and evaluation is as essential. In all medical fields, standards to determine as objectively as possible the benefits of provided care have been proposed according to the principles of evidence-based medicine. It is almost needless to state that this evidence-based approach has made the medical field progress to unprecedented levels. When looking at history, it is interesting to realize that the term evidence-based was only introduced for the first time 30 years ago. The evidence-based principle was an answer to the wide practice variations observed in health care at that time and a manner to use current evidence in making decisions about the care of individual patients.[17] Evidence-based has become since then the watchword in all medical fields as it has been further embedded in medicine over the past decades. However, evidence-based medicine also has its drawbacks, of which we should be aware when evaluating the provided care.

In 2014 Greenhalgh et al. published the first article denunciated the unintended consequences of evidence-based medicine.[18] In 2017, in the Netherlands, the Council for Public Health and Society (*Raad van Volksgezondheid*) came up with a report on the same topic.[19] In those two publications, it was criticized that evidence-based medicine has become nowadays an authority on its own and that, consequently, doctors are often afraid to handle outside the established guidelines. This also results in the fact that the wishes of the patients are often regarded as secondary. Yet, it is important to realize that guidelines are often an over-standardization of care and are not applicable to every individual patient. Another major point of criticism of the evidence-based methodology is that it is primarily driven by statistics and p-values. As a result, there is a tendency to choose numerical outcomes measures such as blood loss, operative time and hospital stay when determining the treatment with the most benefits. This was also the case in **chapter 3** and **chapter 4**. Though, statistical differences in surgical outcomes are not necessarily clinically relevant and of (direct) influence for the patients. While for example from a statistical point of view a difference in blood loss of 50 mL can be relevant, it will probably go unnoticed for the patient. In contrast, a post-operative anemia or the need of a blood transfusion is much more relevant for daily clinical practice.

According to the Dutch report on evidence-based medicine, evidence should be individualized by applying the available evidence in the right context, i.e. *context-based*

medicine. Not only standards should be considered when deciding upon a treatment option but factors related to the patient, the doctor and their environment deserve as much attention. An example of *context-based medicine* was described in **chapter 6**, where we observed that the guideline regarding urinary catheter management after LH was not in line with the opinion of the working floor, i.e. the nurses. A total of 78% of the Dutch hospitals removed the urinary catheter one day after surgery, while 90% of the surveyed nurses believed that direct catheter removed was feasible and 78% would recommend it to a family member or friend. Evidence-based medicine should definitely not be abandoned but other aspects should be additionally considered, even though the statistical support might be less evident. All in all, doctors, nurses and patients should have a more critical approach towards the evidence provided from statistics. Additionally, more attention should be drawn towards relevant outcomes.

In 2006, the concept of *Value-Based Health Care* (VBHC) was introduced by Michael Porter and Elizabeth Olmsted Teisberg.[20] This theory proposes to evaluate medical treatment options based on patient reported outcome measures (PROMs). In that context, we evaluated in **chapter 10** the re-intervention risk after initial therapy and the long-term quality of life of different uterine-sparing techniques, two outcomes that seemed from a patient point of view relevant. Similarly, in **chapter 12**, we evaluated the medical claims filed by patients undergoing a laparoscopic gynecologic procedure. Although we did not directly assess the PROMs, evaluating the claims allowed us to get insight into care judged by patients as being substandard. Understanding the reasons for filing a claim can concomitantly help to improve the quality of care. Also, in **chapter 11**, suggestions were formulated to optimize the postoperative period at home after a minimally invasive procedure. A well-organized postoperative period can be of great value for the patients and also from a financial point of view. In addition, to assure an optimal postoperative recovery and increase patient satisfaction, every aspect of the process should be carefully evaluated. With this in mind, we researched in **chapter 7** the best moment to remove the urinary catheter after LH. Based on the findings of our RCT, immediate catheter removal was recommended as (new) standard practice after LH. Changing catheter policy after LH at a national level may seem a detail in the entire postoperative care process, though it can be of great influence for a patient and her recovery. Although VBHC also has its drawbacks, this concept has given new dimensions to the evaluation of health care and has accelerated the implementation of patient-oriented care.

From the perspectives of VBHC, the financial aspect also plays a crucial role when evaluating the provided care. VBHC aims to maximize patient outcomes for every euro spent. Determining the costs in health care is a complex subject as it is difficult to give

health a financial value. Additionally, data of the actual treatment costs are often not available due to a lack of transparency and differences in agreed prizes between the different hospitals and the health insurance companies. In the discussion comparing VH to TLH (chapter 3), the cost-effectiveness of the procedures is often being brought up as an argument in favor of VH. However, for the Netherlands, no financial data are available on the exact difference between the two surgical approaches. Similarly, in chapter 4, where LESS hysterectomy was compared to conventional LH, data on cost effectiveness were not available either. Although complex political issues might be of influence, additional insights into health care finances in general is absolutely necessary to strengthen the debate regarding the quality of care and even help deciding upon the best treatment.

Conclusion

With this research, we attempted to evaluate clinically relevant aspects of MIS in gynecology and to formulate best practices. For LH, different relevant surgical topics have been covered including an evaluation of the different minimally invasive surgical techniques, the utility of cystoscopy and the best moment to remove the urinary catheter after surgery. These best practices should all together lead to a uniform implementation of LH in the Netherlands. For LM, different topics were evaluated as well, including the potential risks associated with contained tissue extraction and the relative efficacy of LM compared to other uterine-sparing treatment options for fibroids.

Finally, different aspects of MIS were evaluated from patient's perspectives. Outcomes related to patient experience will in our opinion become increasingly important in healthcare in the future. As the MIS technique keeps evolving, we will most certainly face new clinical challenges in the field of gynecology. It is therefore essential to continue monitoring at a national level our procedures based on relevant outcomes only.

Future perspectives

Standardization of care has been proven beneficial from a quality point of view but also for cost reduction. With the development of the guideline for LH a first step towards standardization is made in the field of MIS. With the continuous development of new technology, we would recommend writing similar documents for other surgical procedures, such as LM but also vaginal or abdominal hysterectomy. In addition, an efficient, critical and systematic evaluation of the provided care should take place on a regular basis at a national and local level. In this context, we cannot longer ignore the urgency to systematically collect data for the entire country. Measurement brings knowledge and with the increasing development of information technology (IT), data are easily available. Cohort analyses from large national uniform registration systems should be, at least for surgical innovations, considered as high quality research.

Over the past decades, the introduction of a couple of innovations and techniques in the field of MIS has been inadequate as certain negative outcomes were overlooked. Although not every risk can be anticipated, the introduction of new technology should be more controlled than it has been so far. For a small country as the Netherlands, we would recommend coordinating the implementation of surgical innovations at a national level. Rather than having individual hospitals introducing these new technologies by themselves, expertise should be combined. For this national coordination, we would suggest creating an independent board of experts, who would formulate recommendations regarding the implementation of new technique or technology and perform a PRI that can be shared among hospitals. Meanwhile, surgeons should only be allowed to introduce the innovation in their hospitals in the context of the terms set at national level. One of the main criteria should be that early adopters have the obligation to track outcomes, including outcome measures taking patient's perspectives into consideration. We are aware that creating such a national collaboration is challenging. Yet, we believe that centralizing the introduction of surgical innovations is worth the effort and investment. Unsafe practice will probably be detected earlier and might in the future prevent emotionally driven media attention as we have seen in the past.

In all fields of medicine, challenging clinical topics will always be encountered. To address these issues, the evidence-based approach should remain. However, evidence-based medicine also has limitations that need to be recognized. Specifically for the GRADE assessment, a new system for evaluating the level of evidence of surgical procedures should be developed, and in this new model, we advise a different approach towards the position of cohort studies in relation to randomized controlled trials, since RCTs are not always the best way to answer research questions.

For the practicing doctors, it is important to be continuously aware that the evidence-based approach is primarily based on statistics. As such, it may focus on irrelevant measurable outcomes rather than on outcome measures that are directly relevant for the individual patient. The introduction of VBHC is a revolutionary step in determining valuable care and should be further explored. With the introduction of national databases and the collection of patient specific outcomes, we see great opportunities for providing care that is tailored to a specific patient population.

Finally, the expertise of the doctor and the preferences of the patient should be much more valued that they have been over the last decades. High quality of care can only be provided by shared decision making, i.e. based on a permanent dialogue between the individual patient and his/her doctor supported by evidence from relevant outcome measures. Interestingly, this is exactly in accordance with the first definition of evidence-based medicine formulated thirty years ago.

References

(1) Driessen SR, Baden NL, van Zwet EW, Twijnstra AR, Jansen FW. Trends in the implementation of advanced minimally invasive gynecologic surgical procedures in the Netherlands. J Minim Invasive Gynecol 2015;22:642-647.

(2) Twijnstra AR, Kolkman W, Trimbos-Kemper GC, Jansen FW. Implementation of advanced laparoscopic surgery in gynecology: national overview of trends. J Minim Invasive Gynecol 2010;17:487-492.

(3) Kolkman W, Trimbos-Kemper TC, Jansen FW. Operative laparoscopy in The Netherlands: Diffusion and acceptance. Eur J Obstet Gynecol Reprod Biol 2007;130:245-248.

(4) Guyatt GH, Oxman AD, Vist GE et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336:924-926.

(5) Garry R, Fountain J, Brown J et al. EVALUATE hysterectomy trial: a multicentre randomised trial comparing abdominal, vaginal and laparoscopic methods of hysterectomy. Health Technol Assess 2004;8:1-154.

(6) Aarts JW, Nieboer TE, Johnson N et al. Surgical approach to hysterectomy for benign gynaecological disease. Cochrane Database Syst Rev 2015;CD003677.

(7) Kavanagh BP. The GRADE system for rating clinical guidelines. PLoS Med 2009;6:e1000094.

(8) Ansari MT, Tsertsvadze A, Moher D. Grading quality of evidence and strength of recommendations: a perspective. PLoS Med 2009;6:e1000151.

(9) AAGL Practice Report: Practice guidelines for intraoperative cystoscopy in laparoscopic hysterectomy. J Minim Invasive Gynecol 2012;19:407-411.

(10) Kim SM, Park EK, Jeung IC, Kim CJ, Lee YS. Abdominal, multi-port and single-port total laparoscopic hysterectomy: eleven-year trends comparison of surgical outcomes complications of 936 cases. Arch Gynecol Obstet 2015;291:1313-1319.

(11) https://www.sst.dk/da/udgivelser/2015/nkr-hysterektomi

(12) Deffieux X, Rochambeau B, Chene G et al. Hysterectomy for benign disease: clinical practice guidelines from the French College of Obstetrics and Gynecology. Eur J Obstet Gynecol Reprod Biol 2016;202:83-91.

(13) Neis KJ, Zubke W, Fehr M, Romer T, Tamussino K, Nothacker M. Hysterectomy for Benign Uterine Disease. Dtsch Arztebl Int 2016;113:242-249.

(14) Laparoscopic Uterine Power Morcellation in Hysterectomy and Myomectomy: FDA Safety Communication. [www.fda.gov/medicaldevices/safety/alertsandnotices/ucm393576.htm]. Accessed October 2017

(15) Van Leersum NJ, Snijders HS, Henneman D et al. The Dutch surgical colorectal audit. Eur J Surg Oncol 2013;39:1063-1070.

(16) https://www.igz.nl/onderwerpen/medische-technologie/actuele-onderwerpen/implantatenregister/

(17) Oxman AD, Sackett DL, Guyatt GH. Users' guides to the medical literature. I. How to get started. The Evidence-Based Medicine Working Group. JAMA 1993;270:2093-2095.

(18) Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? BMJ 2014;348:g3725.

(19) https://www.raadrvs.nl/publicaties/item/zonder-context-geen-bewijs

(20) Porter ME. A strategy for health care reform--toward a value-based system. N Engl J Med 2009;361:109-112.