

## Learning curves of open and endoscopic fetal spina bifida closure: a systematic review and meta-analysis

Luc Joyeux<sup>1,2</sup>, Felix De Bie<sup>1</sup>, Enrico Danzer<sup>3</sup>, Francesca M. Russo<sup>1,2,4</sup>, Allan Javaux<sup>5</sup>, Cleisson Fabio Andrioli Peralta<sup>6</sup>, Antonio Afonso Ferreira De Salles<sup>7</sup>, Agnieszka Pastuszka<sup>8</sup>, Anita Olejek<sup>9</sup>, Tim Van Mieghem<sup>10</sup>, Paolo De Coppi<sup>1,11</sup>, Julie Moldenhauer<sup>3</sup>, William E. Whitehead<sup>12</sup>, Michael A. Belfort<sup>13</sup>, Denise Araujo Lapa<sup>14</sup>, Gregorio Lorenzo Acacio<sup>15</sup>, Roland Devlieger<sup>1,4</sup>, Shinjiro Hirose<sup>16</sup>, Diana L. Farmer<sup>16</sup>, Frank Van Calenbergh<sup>17</sup>, N. Scott Adzick<sup>3</sup>, Mark P. Johnson<sup>3</sup>, Jan Deprest<sup>1,2,4,18</sup>

- 1- *Department of Development and Regeneration, Cluster Organ Systems, Biomedical Sciences, Faculty of Medicine, Catholic University of Leuven, Leuven, Belgium*
- 2- *Center for Surgical Technologies, Faculty of Medicine, KU Leuven, Leuven, Belgium*
- 3- *Center for Fetal Diagnosis and Treatment, the Children's Hospital of Philadelphia, and the Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA*
- 4- *Department of Obstetrics and Gynecology, Division Woman And Child, Fetal Medicine Unit, University Hospital of Leuven, Leuven, Belgium*
- 5- *Department of Mechanical Engineering, KU Leuven, Leuven, Belgium*
- 6- *Department of Fetal Medicine, The Heart Hospital, São Paulo, Brazil; Department of Fetal Medicine Pro Matre Hospital, São Paulo, Brazil*
- 7- *Neuroscience Institute, The Heart Hospital, University of São Paulo, São Paulo, Brazil*
- 8- *Department of Descriptive and Topografic Anatomy, Medical University of Silesia in Katowice, School of Medicine with Division of Dentistry, Zabrze, Poland*
- 9- *Department of Gynecology, Obstetrics and Gynecologic Oncology in Bytom, Medical University of Silesia, Bytom, Poland*
- 10- *Department of Obstetrics and Gynecology, Sinai Health System, Mount Sinai Hospital, Toronto, Ontario, Canada*
- 11- *Department of Pediatric Surgery, Great Ormond Street Hospital, University College London Hospitals NHS Foundation trust, London, United Kingdom.*
- 12- *Department of Neurosurgery, Baylor College of Medicine, and Texas Children's Fetal Center, Houston, Texas, USA*
- 13- *Department of Obstetrics and Gynecology, Baylor College of Medicine, and Texas Children's Fetal Center, Houston, Texas, USA*
- 14- *Fetal Therapy Center, Hospital Israelita Albert Einstein, São Paulo, Brazil*
- 15- *Department of Obstetrics and Gynecology, Taubate University, São Paulo, Brazil*
- 16- *Fetal Care and Treatment Center, UC Davis Children's Hospital, Sacramento, California, USA*
- 17- *Department of Neurosurgery, University Hospital of Leuven, Leuven, Belgium*
- 18- *Institute of Women's Health, University College London Hospitals, London, United Kingdom*

### Corresponding author

Jan Deprest, MD PhD

Academic Department Development and Regeneration, Cluster Woman and Child, Biomedical Sciences, Faculty of Medicine, Katholieke Universiteit KU Leuven

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/uog.20389

UZ Herestraat 49, box 1034, 3000 Leuven, Belgium  
E-Mail : [jan.deprest@uzleuven.be](mailto:jan.deprest@uzleuven.be)

**Short title:** Learning curves of fetal surgery for spina bifida

**Keywords** Spina bifida, myelomeningocele, fetal surgery, open fetal surgery, fetoscopy, learning curve, systematic review, meta-analysis.

## ABSTRACT

**Objectives:** The Management Of Myelomeningocele Study (MOMS) trial demonstrated the safety and efficacy of open fetal surgery for spina bifida (SB). Recently developed alternative techniques may reduce maternal risks yet should do without compromising on fetal neuroprotective effects. We aimed to assess the learning curve of different fetal SB closure techniques.

**Methods:** We searched Medline, Web of Science, Embase, Scopus and Cochrane databases and the grey literature to identify relevant articles without language restriction from January 1980 until October 2018. We systematically reviewed and selected studies reporting all consecutive procedures and with a postnatal follow-up  $\geq 12$  months. They also had to report outcome variables necessary to measure the learning curve defined by fetal safety and efficacy. Two independent authors retrieved the data, assessed the quality of the studies and categorized observations into blocks of 30 patients. For meta-analysis, data were pooled using a random-effect model when heterogeneous. To measure the learning curve, we used two complementary methods. With the group splitting method, competency was defined when the procedure provided comparable results to the MOMS trial for 12 outcome variables

Accepted Article

representative for (1) the immediate surgical outcome, (2) short-term neonatal neuroprotection and (3) long-term neuroprotection at  $\geq 12$  months. Then, when the patients' raw data were available, we performed cumulative sum (CUSUM) analysis based on a composite binary outcome defining a successful surgery. It combined four clinically relevant variables for safety (fetal death within 7 days) and for efficacy (neuroprotection at birth).

**Results:** We included 17/6024 (0.3%) studies with low and moderate risks of bias. Fetal SB closure was performed via standard-hysterotomy (n=11), mini-hysterotomy (n=1) or fetoscopy [exteriorized-uterus single-layer (n=1), percutaneous single-layer (n=3) or percutaneous two-layer closure (n=1)]. Only outcomes for the standard-hysterotomy could be meta-analyzed. Overall, outcomes significantly improved with experience. Competency was reached after 35 consecutive cases for standard-hysterotomy and was predicted to be achieved after  $\geq 57$  cases for mini-hysterotomy and  $\geq 56$  for percutaneous two-layer fetoscopy. For percutaneous and uterus-exteriorized single-layer fetoscopy, competency was not respectively reached by cases 81 and 28 available for analysis.

**Conclusions:** The number of cases operated correlates with the outcome of SB fetal closure and ranges from 35 cases for standard-hysterotomy to  $\geq 56-57$  cases for minimally invasive modifications. Our observations provide important information for institutions eager to establish a new fetal center, develop a new technique or train their team, and inform referring clinicians, potential patients and third-parties.

## INTRODUCTION

First described for aircraft manufacturing, the concept of learning curve (LC) is defined in surgery by the acquisition of competency, i.e. the progressive mastering of surgical skills with time and experience required by a surgeon or surgical team to perform a procedure safely and effectively.<sup>1, 2</sup> The LC depends on the patient characteristics, the surgical team, organizational factors such as facilities and equipment and outcomes chosen. Based on its assessment, surgical training programs commonly prescribe a certain number of procedures performed under supervision to certify operators as competent.<sup>1</sup>

Spina bifida aperta (SBA) is a devastating congenital defect, that is progressive in utero. The Management Of Myelomeningocele Study (MOMS) provided level-I evidence that fetal surgery improves outcome as compared to postnatal closure. Prenatal closure reduces the need for ventriculoperitoneal shunting at 12 months and improves neuromotor outcomes at 30 months.<sup>3, 4</sup> The disadvantage of open SBA closure is that it requires maternal laparotomy and a large hysterotomy, both of which put the mother at risk and increase the rate of preterm premature rupture of the membranes (PPROM) and premature delivery. Therefore, alternative less invasive techniques have been recently explored.

Fetoscopic techniques were introduced to overcome these limitations. While the first clinical SBA closures were done fetoscopically, it was abandoned because of its complexity, high fetal mortality and premature delivery.<sup>5</sup> Over the past decade,

several centers have developed a variety of fetoscopic approaches, either with an exteriorized uterus or percutaneously. Ideally a fetoscopic approach should reduce maternal and obstetric risks (PPROM, prematurity, uterine scar dehiscence) without compromising on fetal neuroprotective effects demonstrated for the standard open approach.<sup>3</sup> Innovation is ongoing, yet there is still debate and controversy in terms of clinical implementation, exact technique and current outcomes.<sup>6-8</sup> New centers may be tempted to avoid large hysterotomy and start with a minimally invasive approach. We aimed to determine the nature of the LC of fetal procedures for SBA related to the fetus and not to the mother, irrespective of the approach. This may provide pertinent information regarding surgical technique, safety, efficacy, training methodology and stimulate the development of novel instruments.

## METHODS

### *Study design*

This systematic review and meta-analysis was conducted in accordance with the EQUATOR reporting guidelines, i.e. the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) and the MOOSE (Meta-analysis Of Observational Studies in Epidemiology) guidelines (Table S1).<sup>9-11</sup> We also used the PICO (Patient-Intervention-Comparison-Outcome) framework of interventional studies to form our clinical question and facilitate the literature search.<sup>12</sup> We aimed to measure the LC of different access methods to perform fetal SBA closure by comparing the outcomes of selected parameters to those in the MOMS trial.<sup>3, 4, 13, 14</sup> The MOMS trial is a reference experience using a standardized technique with proven efficacy which was acquired beyond the LC. This systematic review was registered in the PROSPERO registry.<sup>15</sup>

### *Information Sources and Search Technique*

In July 2017, we searched Medline via PubMed, Web of Science, Embase, Scopus and Cochrane databases to identify relevant articles without language restriction published since 1980 (i.e. the beginning of the experimental work around this condition). We also searched the grey literature (100 first hits of Google Scholar) and bibliographies. An update was conducted until October 2018 using the most reliable databases, i.e. PubMed, Embase and Cochrane.<sup>9</sup> Free text and Medical Subject

Accepted Article

Headings (MeSH) terms used and combined for the search were (1) fetal surgery or fetal therapy or fetoscopy or *in utero* surgery or *in utero* therapy or prenatal surgery or prenatal therapy or antenatal surgery or antenatal therapy or intrauterine surgery or intrauterine therapy combined with (2) spina bifida or myeloschisis or myelomeningocele or spinal dysraphism. The term “fetus” being already used in the first term was not used as a third independent term in order not to decrease the number of hits hence miss any paper. Endnote X8.2 (Clarivate Analytics, Boston, MA, USA) was used to eliminate type-I (among different databases) and type-II (in different journals/issues) duplicate reports.<sup>16</sup> Finally, a hand search of the reference lists of all eligible articles was conducted to identify further relevant papers.

#### ***Eligibility Criteria and Study Selection***

To be eligible for inclusion, studies from fetal centers had to report on singleton fetuses with an isolated SBA, who underwent in utero closure and were followed up for at least 12 months. They also had to report outcome variables necessary to measure the LC (as defined below) of consecutive SBA cases. Two authors (L.J., F.D.B) independently reviewed all reports in title-abstract and full-text forms and selected the eligible articles. Any disagreement regarding inclusion of a specific article or interpretation of the data was resolved by discussion and consensus or, if required, by consulting a third author (E.D.). Duplicates, case series of  $\leq 5$  cases, abstracts and conference presentations, letters to the editor and reviews were excluded.

#### ***Outcomes and learning curve assessment***

We used two complementary methods to determine the LC, defined by fetal safety and efficacy, for different surgical techniques. First, the group splitting method was used to define the range of the LC. It consists of chronologically dividing the individual cases into consecutive blocks, which are then analyzed as a group. It is recommended for comparing large numbers, especially when raw data per patient are not available and/or incomplete.<sup>1,2</sup> It precludes determining an exact number of cases required to overcome the LC. Block size was set at 30 cases based on previous studies on the LC for advanced laparoscopic adult digestive<sup>17</sup>, gynecologic<sup>18, 19</sup>, pediatric<sup>20</sup> and fetal<sup>21</sup> surgeries showing a plateau after that number of cases. In our study, competency was defined when the procedure provided similar results as in the MOMS trial for the following clinically relevant and reliable variables that are representative for (1) the immediate surgical outcome (5 variables), (2) short-term neonatal neuroprotection (4 variables) and (3) long-term neuroprotection at the age of one or more (3 variables).<sup>3, 22</sup> Surgical variables were maternal death (from surgery until delivery), postoperative fetal death (within 7 postoperative days), mean operation time (skin-to-skin in minutes), technical failure (aborted or incomplete closure), PPRM <30+0 weeks, preterm delivery <30+0 weeks. Short-term variables were measured at birth and encompassed in utero complete reversal of hindbrain herniation on postnatal MRI, any neonatal treatment (medical or surgical) of a dehiscence or CSF leakage at the closure site, additional surgery at closure site for a dehiscence or CSF leakage, and motor function that improved  $\geq 1$  spinal level.<sup>22</sup> Long-term variables were complete reversal of hindbrain herniation at 12 months, any procedure for CSF diversion (shunt



or endoscopic third ventriculostomy) at 12 months and improved motor function  $\geq 1$  level at 30 months.<sup>3</sup>

Second, when the patients' raw data were available, we used the LC CUmulative SUM (LC-CUSUM) analysis method developed by Biau et al. to precisely determine the LC, i.e. detect or predict how many cases were needed to reach competency.<sup>23</sup>

From that point onwards, skills retention was identified by the Competency CUSUM (C-CUSUM) on the remaining observations.<sup>23</sup> Success of surgery was based on what was defined as first successful fetal surgery for SBA by Adzick et al. relying on four main clinical variables.<sup>24</sup> It was defined as a binary outcome derived from a conjunction of these four reported variables used as binary outcomes.

Two were for safety, i.e. absence of extreme prematurity (delivery  $< 30$  weeks) and of death  $\leq 7$  days from surgery. Two were for efficacy, i.e. evidence of neuroprotection at birth (reversal of hindbrain herniation) and absence of any neonatal treatment (medical or surgical) of a dehiscence or CSF leakage at the closure site.<sup>3, 25</sup> In other words, success was reached when the fetus eventually was delivered beyond 30 weeks, alive, with reversal of hindbrain herniation and without dehiscence or CSF leakage at the closure site. Surgical failure was defined as the need for any neonatal treatment of a dehiscence or CSF leakage. The range (13-30%) of an adequate, i.e. clinically acceptable, failure rate was based on the MOMS trial, the upper limit being very significantly different from the reference lower limit [Delphi decision; 30% (23/77) vs. 13% (10/77) respectively;  $p \leq 0.01$ ].<sup>3</sup> Adequate performance was consequently set when the failure rate was  $\leq 18\%$ , and inadequate performance at

Accepted Article

$\geq 30\%$  with an acceptable deviation (delta) of 5%.<sup>23, 26</sup> The statistical control limit  $h$  was chosen in order to limit the error rates (type I error rate; power 100% when the number of observations tends to infinity).<sup>23</sup> For this purpose, 10,000 simulations were performed at different performance levels and the probability of alarm was computed at various statistical limits (Table S2).<sup>26</sup> As a result, the probability of alarm for the LC-CUSUM at a control limit  $h_{LC}=0.85$  was 99.8%, 74.6% and 23.1% for levels of performance at 18% (adequate), 23% (equivalence) and 30% (inadequate) respectively. In other words, this control limit  $h_{LC}$  was chosen so that the risk to falsely categorize a surgeon as competent was limited to 23.1%, and the risk to falsely categorize a surgeon incompetent was 0.2% (Table S2). Should the threshold  $h_{LC}$  not be reached within the number of reported cases, we performed predictive analytics to forecast when competency would be reached.<sup>27</sup> Our LC-CUSUM algorithm was run with three scenarios: best-case (zero-failure), one-failure and two-failures scenarios. Based upon previous case outcomes (success or failure), each new potential case, except for one or two depending on the scenario, was considered a success until the LC-CUSUM score was just beyond  $h_{LC}$  (Fig.1).

For the C-CUSUM,  $h_C=3$  was chosen so that the risk to falsely categorize the performance as unacceptable was 4.6% and the converse risk to falsely categorize performance as acceptable was 0% (Table S2).

#### ***Data collection and analysis***

A standardized form was used to extract data from the eligible studies (Table S3). Data per center were extracted and categorized into blocks of 30 patients. Categorical

variables were expressed as a percentage and continuous variables as mean and standard deviation. Fetoscopic procedures were categorized based on whether the cannulas were placed following exposure of the uterus through laparotomy (exteriorized-uterus) or through a closed abdomen (percutaneous).

### ***Quality appraisal***

We assessed quality (good, fair and poor) and risk of bias of eligible studies using the adapted criteria outlined in the Cochrane Handbook for Systematic Reviews of Interventions (randomized controlled trial), the Newcastle-Ottawa Scale (NOS) (case-control and cohort studies) and the study quality assessment tool from the American National Institutes of Health (NIH) (case series).<sup>28-30</sup> In cases of attrition (completeness of outcome data) and reporting (selective outcome reporting) bias, corresponding authors were contacted to provide missing outcome data as well as raw chronological data, i.e. the chronological number of the patients and their individual outcomes. When this was unsuccessful, impossible and/or there was missing data for given outcomes, we planned to assess the impact of including those studies in the overall assessment of results using a sensitivity analysis. Finally, we used the GRADE's approach to rate quality of evidence (high, moderate, low or very low) for each outcome per group of fetal surgery technique based on the lowest quality among outcomes.<sup>10, 31</sup>

### ***Meta-analysis***

Methodological and clinical heterogeneity of data per access method were evaluated. Meta-analysis for all outcomes was carried out using MedCalc statistical software

version 15.4 (MedCalc Software bvba, Ostend, Belgium) for PC. Variables were tested for statistical heterogeneity by applying the  $I^2$  statistic test to determine whether data could be pooled by access method.<sup>32</sup> Results were expressed as proportions for categorical variables and mean and standard deviations for continuous variables. Weighted treatment effect was calculated using the fixed or random effects model in case of homogeneity or heterogeneity respectively.<sup>32,33</sup>

### *Statistical analysis*

Date per center and per access method were compared to the MOMS outcomes. Statistical analysis was performed with GraphPad Prism version 7.0d software (GraphPad Inc., La Jolla, CA, USA) for MacOs X. Two-tailed Fisher's exact test was used to compare categorical variables.<sup>34</sup> When we could only obtain mean and standard deviation, we assumed that continuous variables were normally distributed. The unpaired two-tailed independent Student's t-test was used for comparison.<sup>34</sup> When there was no equality of variance, i.e. the 2 groups had a different standard deviation, we added the Welch's correction to the t-test. A p-value <0.05 was considered significant. Finally we applied a LC- and C-CUSUM analysis using a custom-made algorithm in MATLAB® (Mathworks, Natick, MA, USA) based on the model of Biau et al.<sup>23</sup>

## RESULTS

### *Study selection and data extraction*

We identified a total of 6024 publications. The vast majority (n=5939, 98.6%) were excluded based on title and/or abstract (Figure S1). This left 85 (1.4%) studies for full-text evaluation. Subsequently 68 additional studies were excluded, i.e. type II duplicates (n=33), case reports (n=5), or non-consecutive cases (n=5), multicenter experience (n=2) or incomplete perinatal data (n=4). We also excluded 19 studies because the fetal surgery center did not respond [Giessen (Germany, n=2) and one center in Sao Paulo (Brazil, n=5)], was not able to [Nashville, TN (USA, n=4)], Saint Louis, MO (USA, n=2), and Denver-Aurora, CO (USA, n=1)] or was not willing to [Zurich (Switzerland, n=5)] provide unreported outcome measures relevant to our study.

### *Study Characteristics and Risk of Bias*

This left 17 (0.3%) studies for analysis; 11 involving a hysterotomy approach, one via mini-hysterotomy, three via laparotomy and exteriorized-uterus cannulation, and three via percutaneous fetoscopy (Figure S1). Quality assessment revealed 9 studies with low risk, 7 with moderate risk and one with an unclear risk of bias (Table S4). As there were no studies with a high risk of bias, all were included in the meta-analysis.

### *Meta-analysis*

Findings from individual studies were divided into six technically homogenous groups, i.e. based on uterine access and neurosurgical closure technique (Table 1): (1)

via a 6-8 cm hysterotomy and the neurosurgical closure in two to three layers, further referred to as “standard-hysterotomy” (n=11)<sup>3, 4, 13, 14, 35-41</sup>, (2) “mini-hysterotomy”, i.e. same as above yet via a 2.5-3.5 cm hysterotomy and using the operation microscope (n=1)<sup>42</sup>, (3) “percutaneous single-layer” fetoscopic closure (n=3)<sup>43-45</sup>, (4) “percutaneous two-layer” fetoscopic closure (n=1)<sup>46</sup> and (5) “exteriorized-uterus single-layer” fetoscopic closure (n=1)<sup>47</sup>. Only data from the standard-hysterotomy approach (group 1) were meta-analyzed as the required outcome measures were available at different centers (Figure S2).

### ***Quality of evidence of outcomes***

The quality of evidence on each of the safety and efficacy outcome measures in the MOMS was rated high or moderate. The overall quality of evidence for all outcomes was rated high (Table S5). Conversely, the quality of evidence for each outcome measure from all the other (observational) studies was rated low or very low. Also, the overall quality of evidence for these studies was rated very low.

### ***Analysis of outcome data***

All fetal teams were composed of fully-trained anesthesiologists and fetal surgeons, i.e. obstetricians, pediatric neurosurgeons and/or pediatric surgeons. They trained together before opening their center and kept fetal surgeons stable throughout their initial experience.

Overall, outcomes improved with the number of cases operated. Using the group splitting method, we demonstrated that competency for standard-hysterotomy was reached between 31 and 60 cases and remained stable afterwards (Table 2). For mini-

Accepted Article

hysterotomy, competency was not achieved after 45 cases (Table 3). For percutaneous single-layer fetoscopy, despite missing data and based solely on a long operation time, competency was not achieved after 81 cases (Table 4). For percutaneous two-layer fetoscopy, competency was not reached after 47 cases (Table 4). For exteriorized-uterus single-layer fetoscopy, competency was not attained after 28 cases and we could not predict when that would happen due to the limited experience (<30 cases) reported so far (Table 4). Of note, operation times for mini-hysterotomy and all fetoscopic techniques were always significantly longer than for the MOMS standard-hysterotomy (Tables 3 and 4).

Using available raw data, LC- and C-CUSUM analyses helped to determine that competency for standard-hysterotomy performed at the Philadelphia fetal center was precisely reached after 35 cases and remained stable afterwards and before the MOMS trial (Figure 1). In contrast, using our prediction LC-CUSUM model, competency was predicted to occur after at least 57 cases for mini-hysterotomy (Figure 1, Table 3) and after at least 56 cases for percutaneous two-layer fetoscopy (Figure 1, Table 4). We determined that each new failure resulted in four more cases needed to reach competency. In other words, these numbers of cases for mini-hysterotomy were 57, 61 and 65 for the best-case, one-failure and two-failures scenarios respectively. For two-layer percutaneous fetoscopy these numbers were 56, 60 and 64 (Figure 1).

## DISCUSSION

### *Main findings*

This study demonstrates a significant LC for prenatal SBA closure regardless of the approach. The two CUSUM and group splitting methods determined the LC, defined by safety and efficacy, for different surgical techniques and obtained complementary results. 35 (standard-hysterotomy, two to three layers),  $\geq 57$  (mini-hysterotomy; two to three layers) or  $\geq 56$  consecutive cases (percutaneous fetoscopy, two layers) are required to achieve competency for SBA fetal surgery. A minimum of 28 and 81 consecutive cases were needed for exteriorized-uterus and percutaneous single-layer fetoscopy.

### *Clinical interpretation*

The LC is in the order of magnitude of other complex operations. The number of cases needed to achieve competency for standard-hysterotomy was comparable to that of advanced endoscopic procedures such as cholecystectomy<sup>17</sup>, pyloromyotomy<sup>20</sup> or fetoscopic laser coagulation for twin-to-twin transfusion syndrome<sup>21</sup>. For minimally-invasive approaches like mini-hysterotomy and two-layer percutaneous fetoscopy, the numbers were higher and comparable to the LC of laparoscopic colorectal surgery<sup>48, 49</sup> or sacrocolpopexy<sup>18, 19</sup>. The number for single-layer percutaneous fetoscopy was even higher and not comparable. Standard-hysterotomy has the shortest LC and remains the gold-standard technique for safety and efficacy. In contrast, fetoscopy is more demanding, requires additional postnatal procedures and long-term outcomes are still



awaited, however may be neuroprotective and less invasive for the mother.<sup>50</sup> Fetoscopic surgeons should be proficient with advanced endoscopic dissection and suturing. Most fetal centers, apart from one<sup>46</sup>, also believe that surgeons should be able to convert to the standard-hysterotomy approach.<sup>47</sup>

The relevance of operation time as an outcome was demonstrated for standard-hysterotomy as it correlates with complications such as PPROM and preterm delivery.<sup>14, 46</sup> In our study, the group splitting method shows that there is a correlation for all less invasive approaches between unsuccessful surgery and longer operation time. The latter also mean longer anesthesia time, which recently has become a source of concern in children – hence potentially in those unborn.<sup>51</sup> The metabolic effects of the amniodistension CO<sub>2</sub> gas used for all fetoscopic procedures are controversial and being investigated. Whether or not CO<sub>2</sub> insufflation causes fetal acidosis and hypercarbia in human fetuses has not become clear, and the little data that is available suggests that at the most this is minimal.<sup>52</sup> Clearly, in fetal lambs CO<sub>2</sub> insufflation causes acidosis which is operation time-dependent<sup>53-55</sup>, however the different placentation, pressures, uterine wall thickness and vascularity, as well as the fetal physiology do not make the lamb model ideal.

In our study, the numbers obtained are based upon data from fetal centers without any prior formal proctoring<sup>8</sup> or clinical exposure to the procedure studied or to any other surgical technique. New surgical teams should anticipate how, and over what time, they can realistically reach competency so as to train their team in offering the procedure safely and effectively, and retaining their skills thereafter. This may be

Accepted Article

difficult for a condition as rare as SBA, narrow inclusion criteria and limited uptake of fetal surgery by patients.<sup>56</sup> Preclinical steps may include proctoring, in-house and exported training<sup>56</sup>, simulation, and/or other methods that shorten the LC.<sup>1, 2, 8</sup> Clinically, our study confirms, and expands upon, the previous position statement that “a new fetal SBA repair team will be considered experienced once they have performed or participated in at least 5 cases”.<sup>7, 8</sup> We would recommend on-site and off-site supervision from expert centers until competency has been achieved and proven, using our LC-CUSUM assessment method. On-site supervision could be defined by at least one expert joining the new team at the operating table for  $\geq 5$  cases. Off-site supervision could mean further supervision during the LC phase using pre-, intra- and/or postoperative tele-mentoring via phone or video conference calls.

#### ***Strengths and limitations***

The strengths of our study are that we adhered to international guidelines and we included only studies with low or moderate risk of bias. Criteria defining successful fetal closure were based both on fetal safety and efficacy, as they are both clinically relevant outcome measures. Furthermore, the two valid and complementary group splitting and CUSUM methods obtained similar results and their combination allowed us to precisely determine the LC.

We are also aware of its limitations. First, this study was not designed to assess the maternal effects despite their importance. It would have required the measurement of a specific LC of prenatal surgery related to the mother. The latter would need higher numbers of patients, as maternal complications are rarer, and long-term maternal

obstetric and reproductive outcomes which have not yet been comprehensively reported.<sup>57, 58</sup> Second, we could not assess the LC of individual surgeons but rather of fetal teams. This surgery is indeed performed in tertiary fetal centers with a multidisciplinary team.<sup>8</sup> Third, inclusion criteria for fetal surgery were not identical to those of the MOMS, although we expected the majority of cases would comply. This shortcoming was taken into account by the clinical and statistical heterogeneity analysis. Fourth, the LC of mini-hysterotomy and two-layer percutaneous fetoscopy were forecasted based on predictive analytics using three scenarios from a mathematical modeling for the LC-CUSUM analysis. These have not been confirmed yet due to the limited number of cases. Fifth, we could not acquire the raw data from all fetal centers, because they were not always able or willing to share it, which precluded us from using the CUSUM analysis throughout. Therefore, the LC of single-layer percutaneous and exteriorized-uterus fetoscopies were solely measured using the group splitting method which might lead to an unfair comparison to the MOMS trial standard-hysterotomy. Sixth, this review compares experiences with different in-utero neurosurgical SBA closure techniques, which are currently a matter of debate and continue to evolve. Seventh, even a comparison of identical techniques between different centers is challenging. Pioneering centers from the MOMS trial had already years of experience with experimental and clinical fetal surgeries for SBA and other indications and others may not have comparable experience.<sup>59</sup> Finally, combining standard-hysterotomy data in a meta-analysis increased the overall

variability, in turn reducing power.<sup>60</sup> We countered this by using a random effect model to adjust for heterogeneity across studies.

## CONCLUSIONS

The number of cases needed for completing the LC for standard-hysterotomy is the smallest ( $n=35$ ) among all approaches. This technique remains the gold-standard technique for safe and effective fetal SBA closure as demonstrated in the MOMS trial. We also show, using mathematical modelling, that the number is higher for mini-hysterotomy ( $n\geq 57$ ) and percutaneous fetoscopy ( $n\geq 56$ ). Our observations provide important information for institutions eager to establish a new fetal center, develop a new technique or train their team, and inform referring clinicians, potential patients and third-parties.

## **ACKNOWLEDGMENTS**

We are very grateful to Dr Alicja K. Olejnik (Politechnika Łódzka, Międzyresortowy Instytut Techniki Radiacyjnej, Poland) for translating one article from Polish and to Prof Guillermo Antinolo (Virgen de Rocio University Hospital of Sevilla, Spain) for sharing data of his 3 cases. We are also indebted to Lindsay Waqar from the Children's Hospital of Philadelphia and Sharan Johal from UC Davis Health, research coordinators, for their help in collecting data.

## **FUNDING**

LJ is funded through an Innovative Engineering for Health award by the Wellcome Trust (WT101957) and the Engineering and Physical Sciences Research Council (EPSRC) (NS/A000027/1). His research is part of the Guided Instrumentation for Fetal Therapy and Surgery (GIFT-Surg) project. JD is supported by the Great Ormond Street Hospital Charity fund.

## **CONFLICT OF INTEREST** <sup>[L]</sup><sub>[SEP]</sub>

The authors declare no competing financial interests.

## **SYSTEMATIC REVIEW REGISTRATION NUMBER**

CRD42016033285 (PROSPERO, February 16th, 2016).

## REFERENCES

1. Harrysson IJ, Cook J, Sirimanna P, Feldman LS, Darzi A, Aggarwal R. Systematic review of learning curves for minimally invasive abdominal surgery: a review of the methodology of data collection, depiction of outcomes, and statistical analysis. *Ann Surg* 2014; **260**: 37-45.
2. Khan N, Abboudi H, Khan MS, Dasgupta P, Ahmed K. Measuring the surgical 'learning curve': methods, variables and competency. *BJU Int* 2014; **113**: 504-508.
3. Adzick NS, Thom EA, Spong CY, Brock JW, 3rd, Burrows PK, Johnson MP, Howell LJ, Farrell JA, Dabrowiak ME, Sutton LN, Gupta N, Tulipan NB, D'Alton ME, Farmer DL, Investigators M. A randomized trial of prenatal versus postnatal repair of myelomeningocele. *N Engl J Med* 2011; **364**: 993-1004.
4. Farmer DL, Thom EA, Brock JW, 3rd, Burrows PK, Johnson MP, Howell LJ, Farrell JA, Gupta N, Adzick NS, Management of Myelomeningocele Study I. The Management of Myelomeningocele Study: full cohort 30-month pediatric outcomes. *Am J Obstet Gynecol* 2018; **218**: 256 e251-256 e213.
5. Bruner JP, Tulipan NB, Richards WO, Walsh WF, Boehm FH, Vrabcak EK. In utero repair of myelomeningocele: a comparison of endoscopy and hysterotomy. *Fetal Diagn Ther* 2000; **15**: 83-88.
6. Flake A. Percutaneous minimal-access fetoscopic surgery for myelomeningocele - not so minimal! *Ultrasound Obstet Gynecol* 2014; **44**: 499-500.
7. Belfort M, Deprest J, Hecher K. Current controversies in prenatal diagnosis 1: in utero therapy for spina bifida is ready for endoscopic repair. *Prenat Diagn* 2016; **36**: 1161-1166.
8. Cohen AR, Couto J, Cummings JJ, Johnson A, Joseph G, Kaufman BA, Litman RS, Menard MK, Moldenhauer JS, Pringle KC, Schwartz MZ, Walker WO, Jr., Warf BC, Wax JR, Force MMCM-FMT. Position statement on fetal myelomeningocele repair. *Am J Obstet Gynecol* 2014; **210**: 107-111.
9. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; **6**: e1000097.
10. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P, Schunemann HJ. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011; **64**: 383-394.
11. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; **283**: 2008-2012.
12. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. *ACP J Club* 1995; **123**: A12-13.
13. Tulipan N, Wellons JC, 3rd, Thom EA, Gupta N, Sutton LN, Burrows PK, Farmer D, Walsh W, Johnson MP, Rand L, Tolivaisa S, D'Alton M E, Adzick NS, Investigators M. Prenatal surgery for myelomeningocele and the need for cerebrospinal fluid shunt placement. *J Neurosurg Pediatr* 2015; **16**: 613-620.
14. Johnson MP, Bennett KA, Rand L, Burrows PK, Thom EA, Howell LJ, Farrell JA, Dabrowiak ME, Brock JW, 3rd, Farmer DL, Adzick NS, Management of Myelomeningocele Study I. The Management of Myelomeningocele Study: obstetrical outcomes and risk factors for obstetrical complications following prenatal surgery. *Am J Obstet Gynecol* 2016; **215**: 778 e771-778 e779.
15. Joyeux L. There is a different learning curve for open and fetoscopic spina bifida surgery – a systematic review and meta-analysis. In *Prospero*, vol PROSPERO, International prospective register of systematic reviews (Feb 16, 2016 edn). Centre for Reviews and Dissemination (CRD), University of York  
NHS National Institute for Health Research: York, UK, 2016.
16. Qi X, Yang M, Ren W, Jia J, Wang J, Han G, Fan D. Find duplicates among the PubMed, EMBASE, and Cochrane Library Databases in systematic review. *PLoS One* 2013; **8**: e71838.
17. Moore MJ, Bennett CL. The learning curve for laparoscopic cholecystectomy. The Southern Surgeons Club. *Am J Surg* 1995; **170**: 55-59.

18. Claerhout F, Verguts J, Werbrouck E, Veldman J, Lewi P, Deprest J. Analysis of the learning process for laparoscopic sacrocolpopexy: identification of challenging steps. *Int Urogynecol J* 2014; **25**: 1185-1191.
19. Vossaert K, Housmans S, Pacquee S, Callewaert G, Cattani L, Van der Aa F, Wolthuis A, D'hoore A, Roelandt P, Deprest J. Laparoscopic sacrocolpopexy is as safe in septuagenarians or elder as in younger women. *Gynecol Surg* 2018; **15**: 11.
20. Oomen MW, Hoekstra LT, Bakx R, Heij HA. Learning curves for pediatric laparoscopy: how many operations are enough? The Amsterdam experience with laparoscopic pyloromyotomy. *Surg Endosc* 2010; **24**: 1829-1833.
21. Peeters SH, Van Zwet EW, Oepkes D, Lopriore E, Klumper FJ, Middeldorp JM. Learning curve for fetoscopic laser surgery using cumulative sum analysis. *Acta Obstet Gynecol Scand* 2014; **93**: 705-711.
22. Johnson MP, Sutton LN, Rintoul N, Crombleholme TM, Flake AW, Howell LJ, Hedrick HL, Wilson RD, Adzick NS. Fetal myelomeningocele repair: short-term clinical outcomes. *Am J Obstet Gynecol* 2003; **189**: 482-487.
23. Biau DJ, Porcher R. A method for monitoring a process from an out of control to an in control state: Application to the learning curve. *Stat Med* 2010; **29**: 1900-1909.
24. Adzick NS, Sutton LN, Crombleholme TM, Flake AW. Successful fetal surgery for spina bifida. *Lancet* 1998; **352**: 1675-1676.
25. Kabagambe SK, Jensen GW, Chen YJ, Vanover MA, Farmer DL. Fetal Surgery for Myelomeningocele: A Systematic Review and Meta-Analysis of Outcomes in Fetoscopic versus Open Repair. *Fetal Diagn Ther* 2018; **43**: 161-174.
26. Papanna R, Biau DJ, Mann LK, Johnson A, Moise KJ, Jr. Use of the Learning Curve-Cumulative Summation test for quantitative and individualized assessment of competency of a surgical procedure in obstetrics and gynecology: fetoscopic laser ablation as a model. *Am J Obstet Gynecol* 2011; **204**: 218 e211-219.
27. Song X, Mitnitski A, Cox J, Rockwood K. Comparison of machine learning techniques with classical statistical models in predicting health outcomes. *Stud Health Technol Inform* 2004; **107**: 736-740.
28. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JA, Cochrane Bias Methods G, Cochrane Statistical Methods G. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011; **343**: d5928.
29. National Institutes of Health N. Study Quality Assessment Tools. <http://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools.>
30. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. In *Clin Epidemiol*, vol 2008 Wells GA (ed). Ottawa Hospital Research Institute: Ottawa, Canada, 2008.
31. Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S, Guyatt GH. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011; **64**: 401-406.
32. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557-560.
33. Borenstein M, Hedges LV, Higgins J, Rothstein HR. *Introduction to meta-analysis*. Wiley Online Library: Chichester, UK, 2009.
34. Mehanna H, Al-Maqbili T, Carter B, Martin E, Campain N, Watkinson J, McCabe C, Boelaert K, Franklyn JA. Differences in the recurrence and mortality outcomes rates of incidental and nonincidental papillary thyroid microcarcinoma: a systematic review and meta-analysis of 21 329 person-years of follow-up. *J Clin Endocrinol Metab* 2014; **99**: 2834-2843.
35. Zamlyniski J, Olejek A, Bohosiewicz J, Bodzek P, Manka G, Grettka K, Paliga M, Gajewska A. [Perinatal results of intrauterine open fetal surgery of fetuses diagnosed with myelomeningocele--the clinical report of ten cases]. *Ginekol Pol* 2007; **78**: 647-651.
36. Moldenhauer JS, Soni S, Rintoul NE, Spinner SS, Khalek N, Martinez-Poyer J, Flake AW, Hedrick HL, Peranteau WH, Rendon N, Koh J, Howell LJ, Heuer GG, Sutton LN, Johnson MP, Adzick



NS. Fetal myelomeningocele repair: the post-MOMS experience at the Children's Hospital of Philadelphia. *Fetal Diagn Ther* 2015; **37**: 235-240.

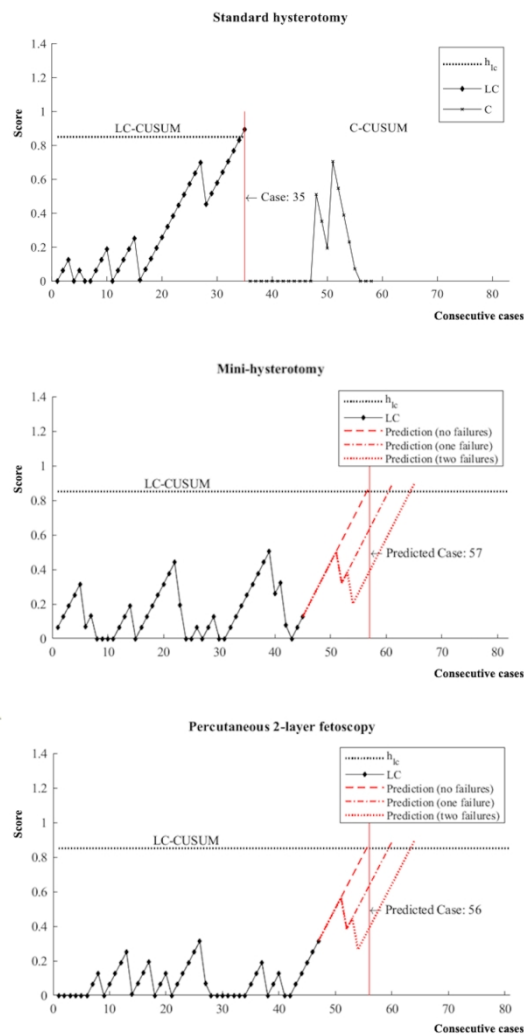
37. Hisaba WJ, Cavalheiro S, Almodim CG, Borges CP, de Faria TC, Araujo Junior E, Nardozza LM, Moron AF. Intrauterine myelomeningocele repair postnatal results and follow-up at 3.5 years of age--initial experience from a single reference service in Brazil. *Childs Nerv Syst* 2012; **28**: 461-467.
38. Farmer DL, von Koch CS, Peacock WJ, Danielpour M, Gupta N, Lee H, Harrison MR. In utero repair of myelomeningocele: experimental pathophysiology, initial clinical experience, and outcomes. *Arch Surg* 2003; **138**: 872-878.
39. Danzer E, Gerdes M, Bebbington MW, Zarnow DM, Adzick NS, Johnson MP. Preschool neurodevelopmental outcome of children following fetal myelomeningocele closure. *Am J Obstet Gynecol* 2010; **202**: 450 e451-459.
40. Bruner JP, Tulipan N, Paschall RL, Boehm FH, Walsh WF, Silva SR, Hernanz-Schulman M, Lowe LH, Reed GW. Fetal surgery for myelomeningocele and the incidence of shunt-dependent hydrocephalus. *JAMA* 1999; **282**: 1819-1825.
41. Bennett KA, Carroll MA, Shannon CN, Braun SA, Dabrowiak ME, Crum AK, Paschall RL, Kavanaugh-McHugh AL, Wellons JC, 3rd, Tulipan NB. Reducing perinatal complications and preterm delivery for patients undergoing in utero closure of fetal myelomeningocele: further modifications to the multidisciplinary surgical technique. *J Neurosurg Pediatr* 2014; **14**: 108-114.
42. Botelho RD, Imada V, Rodrigues da Costa KJ, Watanabe LC, Rossi Junior R, De Salles AAF, Romano E, Peralta CFA. Fetal Myelomeningocele Repair through a Mini-Hysterotomy. *Fetal Diagn Ther* 2017; **42**: 28-34.
43. Verbeek RJ, Heep A, Maurits NM, Cremer R, Hoving EW, Brouwer OF, van der Hoeven JH, Sival DA. Fetal endoscopic myelomeningocele closure preserves segmental neurological function. *Dev Med Child Neurol* 2012; **54**: 15-22.
44. Degenhardt J, Schurg R, Winarno A, Oehmke F, Khaleeva A, Kawecki A, Enzensberger C, Tinneberg HR, Faas D, Ehrhardt H, Axt-Flidner R, Kohl T. Percutaneous minimal-access fetoscopic surgery for spina bifida aperta. Part II: maternal management and outcome. *Ultrasound Obstet Gynecol* 2014; **44**: 525-531.
45. Kohl T. Percutaneous minimally invasive fetoscopic surgery for spina bifida aperta. Part I: surgical technique and perioperative outcome. *Ultrasound Obstet Gynecol* 2014; **44**: 515-524.
46. Lapa Pedreira DA, Acacio GL, Goncalves RT, Sa RAM, Brandt RA, Chmait RH, Kontopoulos EV, Quintero RA. Percutaneous fetoscopic closure of large open spina bifida using a bilaminar skin substitute. *Ultrasound Obstet Gynecol* 2018; **52**: 458-466.
47. Belfort MA, Whitehead WE, Shamshirsaz AA, Bateni ZH, Olutoye OO, Olutoye OA, Mann DG, Espinoza J, Williams E, Lee TC, Keswani SG, Ayres N, Cassady CI, Mehollin-Ray AR, Sanz Cortes M, Carreras E, Peiro JL, Ruano R, Cass DL. Fetoscopic Open Neural Tube Defect Repair: Development and Refinement of a Two-Port, Carbon Dioxide Insufflation Technique. *Obstet Gynecol* 2017; **129**: 734-743.
48. Jamali FR, Soweid AM, Dimassi H, Bailey C, Leroy J, Marescaux J. Evaluating the degree of difficulty of laparoscopic colorectal surgery. *Arch Surg* 2008; **143**: 762-767; discussion 768.
49. Tekkis PP, Senagore AJ, Delaney CP, Fazio VW. Evaluation of the learning curve in laparoscopic colorectal surgery: comparison of right-sided and left-sided resections. *Ann Surg* 2005; **242**: 83-91.
50. Joyeux L, Engels AC, Russo FM, Jimenez J, Van Mieghem T, De Coppi P, Van Calenbergh F, Deprest J. Fetoscopic versus Open Repair for Spina Bifida Aperta: A Systematic Review of Outcomes. *Fetal Diagn Ther* 2016; **39**: 161-171.
51. FDA US. FDA review results in new warnings about using general anesthetics and sedation drugs in young children and pregnant women. *FDA drug safety communication* 2017.
52. Baschat AA, Ahn ES, Murphy J, Miller JL. Fetal blood gas values during fetoscopic myelomeningocele repair performed under carbon dioxide insufflation. *Ultrasound Obstet Gynecol* 2018. DOI:10.1002/uog.19083.



53. Gratacos E, Wu J, Devlieger R, Van de Velde M, Deprest JA. Effects of amniodistention with carbon dioxide on fetal acid-base status during fetoscopic surgery in a sheep model. *Surg Endosc* 2001; **15**: 368-372.
54. Skinner S, Crossley K, Amberg B, Kashyap A, Hooper S, Deprest JA, Hodges R, DeKoninck P. The effects of partial amniotic carbon dioxide insufflation in an ovine model. *Prenat Diagn* 2018; **38**: 994-1003.
55. Sanz Cortes M, Castro E, Sharhan D, Torres P, Yopez M, Espinoza J, Shamshirsaz AA, Nassr AA, Poppek E, Whitehead W, Belfort MA. Amniotic membrane and placental histopathological findings after open and fetoscopic prenatal neural tube defect repair. *Prenat Diagn* 2019; **39**: 269-279.
56. Ovaere C, Eggink A, Richter J, Cohen-Overbeek TE, Van Calenbergh F, Jansen K, Oepkes D, Devlieger R, De Catte L, Deprest JA. Prenatal diagnosis and patient preferences in patients with neural tube defects around the advent of fetal surgery in Belgium and Holland. *Fetal Diagn Ther* 2015; **37**: 226-234.
57. Wilson RD, Johnson MP, Flake AW, Crombleholme TM, Hedrick HL, Wilson J, Adzick NS. Reproductive outcomes after pregnancy complicated by maternal-fetal surgery. *Am J Obstet Gynecol* 2004; **191**: 1430-1436.
58. Wilson RD, Lemerand K, Johnson MP, Flake AW, Bebbington M, Hedrick HL, Adzick NS. Reproductive outcomes in subsequent pregnancies after a pregnancy complicated by open maternal-fetal surgery (1996-2007). *Am J Obstet Gynecol* 2010; **203**: 209 e201-206.
59. Joyeux L, De Bie F, Danzer E, Van Mieghem T, Flake AW, Deprest J. Safety and efficacy of fetal surgery techniques to close a spina bifida defect in the fetal lamb model: A systematic review. *Prenat Diagn* 2018; **38**: 231-242.
60. Bangdiwala SI, Bhargava A, O'Connor DP, Robinson TN, Michie S, Murray DM, Stevens J, Belle SH, Templin TN, Pratt CA. Statistical methodologies to pool across multiple intervention studies. *Transl Behav Med* 2016; **6**: 228-235.

## FIGURES LEGENDS

**Figure 1 – Learning curve analysis of fetal surgery techniques using cumulative sum analysis (CUSUM).** We used data from consecutive cases of 3 fetal centers: Philadelphia (PA, USA) for standard-hysterotomy, San Paolo (Brazil) center for mini-hysterotomy and another San Paolo center for two-layer percutaneous fetoscopy. When the threshold  $h_{LC}$  was not reached within the number of reported cases, we performed predictive analytics to forecast when competency would be reached. Our LC-CUSUM algorithm was run with three scenarios: best-case (zero-failure), one-failure and two-failures scenarios. Each new potential case, except for one or two, was considered a success until the LC-CUSUM score was just beyond  $h_{LC}$ . Abbreviations: LC, learning curve; C, competency.



**Table 1 - Clinically homogenous groups based on uterine access and neurosurgical technique.** Closure of the defect was categorized as either being a “repair,” i.e., with prior dissection around the defect followed by skin repair or grafting, or “coverage” when the defect was covered yet without prior dissection of the placode. Abbreviations: MOMS, management of myelomeningocele study; NA, non-applicable; NS, non-specified.

Approaches for fetal surgery closure	Standard-hysterotomy	Mini-hysterotomy	Percutaneous single-layer fetoscopy	Percutaneous two-layer fetoscopy	Exteriorized-uterus single layer fetoscopy
<b>Studies</b>	n=11 <sup>3, 4, 13, 14, 35-41</sup>	n=1 <sup>42</sup>	n=3 <sup>43-45</sup>	n=1 <sup>46</sup>	n=1 <sup>47</sup>
<b># of cases</b>	347 (91 MOMS, 256 others)	45	81	47	28
<b>Centers</b>	Multicenters	Sao Paulo, Brazil	Bonn-Giessen, Germany	Sao Paulo, Brazil	Houston, TX, USA
<b>Laparotomy</b>	Large ≥15cm Classic	Large ≥15cm Classic	Small 1cm Seldinger technique	Small 1cm Seldinger technique	Large ≥15cm Classic
<b>Hysterotomy</b>	One 6-8 cm	One 2.5-3.5 cm	Three to four 3-4x0.5 cm	Three to four 0.8, 3x0.4 cm	Two 2x0.5 cm
<b>Uterine access</b>	Incision	Incision	Seldinger technique	Seldinger technique	Seldinger technique
<b>Size of cannulas (internal diameter)</b>	NA	NA	4mm (12 French)	One of 5mm Two/three of 3.7 mm (11 French)	4mm (12 French)
<b>Neurosurgical closure technique</b>	Repair	Repair	Repair	Repair	Repair
<b>Placode dissection</b>	Yes	Yes	Yes	Yes	Yes
<b>Complete placode untethering</b>	Yes	NS	NS	NS	NS
<b>Layers over the placode</b>	2-3 layers - Dura - Myofascia - Skin (or Alloderm®)	2-3 layers - Arachnoid - Dura - Skin	1 layer - Skin patch (Gore Preclude® or Surgisis®)	2 layers - Fascia patch (biocellulose Bionext®) - Skin (or Integra® Dermal Regeneration Template)	1 layer - unified layer (skin and dura)

**Table 2 - Comparison of results from the standard-hysterotomy technique used by various fetal centers to the results of the MOMS trial using the group splitting method.** Data are expressed in percentages (ratios) or mean  $\pm$  standard deviation. An asterisk \* indicates a significantly worse result (p value <0.5). Abbreviations: MOMS, management of myelomeningocele study; PPROM, preterm premature rupture of membrane; HH, hindbrain herniation; CSF, cerebrospinal fluid; NS, non-specified.

Comparison of standard-hysterotomy centers v. MOMS trial	Reference MOMS trial	5 meta-analyzed centers <sup>35, 37-40</sup>	Danzer 2010 <sup>39</sup>	3 centers from MOMS trial <sup>3, 4, 13, 14</sup>	2 meta-analyzed centers <sup>36, 41</sup>
Fetal centers	3 MOMS centers	Vanderbilt, Philadelphia, San Francisco, Sao Paulo, Katowice-Bytom	Philadelphia	3 MOMS centers	Vanderbilt, Philadelphia,
Total number of patients included for analysis	91	85	28	30	143
Block of experience		Block I	Block II	Block III	Block IV
<b>Surgical outcomes</b>					
Maternal death	0.0% (0/91)	0.0% (0/85)	0.0% (0/28)	0.0% (0/91)	0.0% (0/143)
Postoperative death $\leq 7d$	2.2% (2/91)	3.4% (3/85)	0.0% (0/28)	2.2% (2/91)	2.7% (4/142)
Mean operation time (min)	105 $\pm$ 22	NS	NS	105 $\pm$ 22	78 $\pm$ 12
Technical failure	0.0% (0/91)	1.3% (1/85)	0.0% (0/28)	0.0% (0/91)	1.1% (2/143)
PPROM <30+0 weeks gestation	NS	18.9% (8/40)	3.6% (1/28)	NS	16.7% (16/96)
Delivery <30+0 weeks	11.0% (10/91)	22.5% (19/85)	7.1% (2/28)	11.0% (10/91)	8.4% (12/139)
<b>Short-term neonatal neuroprotection</b>					
In utero complete reversal of HH	NS	41.4% (23/56)	40.8% (11/27)	NS	71.1% (59/83)
Any treatment at repair site	13.0% (10/77)	3.7% (1/27)	0.0% (0/27)	13.0% (10/77)	5.4% (7/124)
Additional recoverage at repair site	2.6% (2/77)	17.4% (6/35) *	0.0% (0/27)	2.6% (2/77)	3.1% (4/124)
Improved motor function	NS	59.2% (25/42)	59.3% (16/27)	NS	55.0% (44/80)
<b>Long-term neuroprotection</b>					
Complete reversal of HH at 12 months	35.7% (25/70)	NS	NS	35.7% (25/70)	NS
CSF diversion at 12 months	44.0% (40/91)	45.3% (18/40)	48.1% (13/27)	44.0% (40/91)	35.3% (45/128)
Improved motor function at 30 months	37.9% (33/87)	52.9% (9/17)	52.9% (9/17)	37.9% (33/87)	NS

**Table 3 – Comparison of results from the Brazilian mini-hysterotomy technique to the results of the MOMS trial using the group splitting method.** Data are expressed in percentages (ratios) or mean  $\pm$  standard deviation. An asterisk \* indicates a significantly worse result ( $p$  value  $<0.5$ ). Abbreviations: OFS, open fetal surgery; MOMS, management of myelomeningocele study; PPROM, preterm premature rupture of membrane; HH, hindbrain herniation; CSF, cerebrospinal fluid; NS, non-specified; NA, non-applicable.

Comparison of mini-hysterotomy vs. MOMS trial	Reference MOMS trial	Botelho 2017 <sup>42</sup>		Statistical analysis
		30 Block I	15 Block II	
Number of patients Block of experience	91	30 Block I	15 Block II	
<b>Surgical outcomes</b>				
Maternal death	0.0% (0/91)	0.0% (0/30)	0.0% (0/15)	Fisher exact test
Postoperative death $\leq 7$ d	2.2% (2/91)	0.0% (0/30)	0.0% (0/15)	Fisher exact test
Mean operation time (min)	105 $\pm$ 22	225 $\pm$ 38 *	211 $\pm$ 41 *	t-test with Welch correction
Technical failure	0.0% (0/91)	0.0% (0/30)	0.0% (0/15)	Fisher exact test
PPROM $<30+0$ weeks	NS	0.0% (0/30)	6.7% (1/15)	NA
Delivery $<30+0$ weeks	11.0% (10/91)	6.7% (2/30)	0.0% (0/15)	Fisher exact test
<b>Short-term neonatal neuroprotection</b>				
In utero complete reversal of HH	NS	36.7% (11/30)	20.0% (3/15)	NA
Any treatment at repair site	13.0% (10/77)	10.0% (3/30)	6.7% (1/15)	NA
Additional recoverage at repair site	2.6% (2/77)	10.0% (3/30)	6.7% (1/15)	Fisher exact test
Improved motor function	NS	43.3% (13/30)	20.0% (3/15)	NA
<b>Short-term neuroprotection</b>				
Complete reversal of HH at 12 months	35.7% (25/70)	36.7% (11/30)	20.0% (3/15)	Fisher exact test
CSF diversion at 12 months	44.0% (40/91)	30.0% (10/30)	57.1% (8/14)	Fisher exact test
Improved motor function at 30 months	37.9% (33/87)	43.3% (13/30)	NS	Fisher exact test

**Table 4 – Comparison of results from fetoscopy centers to the results of the MOMS trial using the group splitting method.** Data are expressed in percentages (ratios) or mean ± standard deviation. An asterisk \* indicates a significantly worse result and # a significantly better result (p value <0.5). Abbreviations: OFS, open fetal surgery; MOMS, management of myelomeningocele study; H, hysterotomy; PPROM, preterm premature rupture of membrane; HH, hindbrain herniation; CSF, cerebrospinal fluid; NS, non-specified.

Comparison of fetoscopy vs MOMS trial	Reference MOMS trial	Verbeek 2012 <sup>43</sup>	Degenhardt 2014 & Kohl 2014 <sup>44, 45</sup>			Lapa-Pedreira 2018 <sup>46</sup>		Belfort 2017 <sup>47</sup>
Technique	Standard-H	Percutaneous 3/4-port 1-layer			Percutaneous 3/4-port 2-layer		Exteriorized-uterus 2-port	
Number of patients	91	19	30	21	30	17	28	
Block of experience		Block I	Block II	Block III	Block I	Block II	Block I	
<b>Surgical outcomes</b>								
Maternal death	0.0% (0/91)	0.0% (0/19)	0.0% (0/30)	0.0% (0/21)	0.0% (0/30)	0.0% (0/17)	0.0% (0/28)	
Postoperative death ≤7d	2.2% (2/91)	15.8% (3/19) *	3.3% (1/30)	0.0% (0/21)	3.3% (1/30)	0.0% (0/17)	0.0% (0/28)	
Mean operation time (min)	105 ±22	NS	241 ±34 *	196 ±34 *	220 ±70 *	148 ±75 *	264 ±76 *	
Technical failure	0.0% (0/91)	15.8% (3/19) *	0.0% (0/30)	4.8% (1/21)	6.7% (2/30)	0.0% (0/17)	21.4% (6/28) *	
PPROM <30+0 weeks	NS	NS	46.7% (14/30)	65.0% (13/20)	21.4% (6/28)	50.0% (9/18)	0% (0/22)	
Delivery <30+0 weeks	11.0% (10/91)	NS	13.3% (4/30)	10.0% (2/20)	14.3% (4/28)	17.6% (3/17)	4.5% (1/22)	
<b>Short-term neonatal neuroprotection</b>								
In utero complete reversal of HH	NS	15.4% (2/13)	NS	NS	56.5% (13/23)	53.3% (8/15)	57.1% (12/21)	
Any treatment at repair site	13.0% (10/77)	NS	NS	NS	28.6% (8/28)	17.6% (3/17)	36.4% (8/22) *	
Additional coverage at repair site	2.6% (2/77)	NS	NS	NS	17.9% (5/28) *	11.8% (2/17)	9.1% (2/22)	
Improved motor function	NS	69.2% (9/13)	NS	NS	69.2% (18/26)	33.3% (5/15)	72.7% (16/22)	
<b>Long-term neuroprotection</b>								
Complete reversal of HH at 12 months	35.7% (25/70)	0% (0/7)	NS	NS	75.0% (9/12) #	NS	54.5% (12/22)	
CSF diversion at 12 months	44.0% (40/91)	30.8% (4/13)	NS	NS	69.2% (18/26) *	50.0% (6/12)	42.9% (9/21)	

<b>Improved motor function at 30 months</b>	<b>37.9% (33/87)</b>	NS	NS	NS	NS	NS	<b>72.7% (16/22) #</b>
---	--------------------------	----	----	----	----	----	----------------------------