The amniotic fluid index and oligohydramnios: a deeper dive into the shallow end

Joseph R. Wax, MD; Michael G. Pinette, MD

Clinical Guidance

The rationale for this recommendation is based on the findings of randomized clinical trials (RCTs) comparing the SDP to the AFI in evaluating AFV in singleton pregnancies. Taken individually and collectively, these studies uniformly demonstrate similar perinatal outcomes with both techniques, whereas the AFI incurs increased rates of diagnosing oligohydramnios and obstetrical interventions.12–19 Our review of the same evidence leads us to draw very different conclusions and suggest revising existing clinical recommendations. It is our clinical opinion that the AFI should be an acceptable option for AFV assessment, including diagnosing oligohydramnios. In addition, we contend that the very performance characteristics leading to recommending SDP vs AFI in diagnosing oligohydramnios may support the preferential use of AFI vs SDP in this regard.

Randomized Clinical Trials and Clinical Guidance

Of note, 6 published RCTs compare the AFI and the SDP as screening tests for preventing adverse pregnancy outcomes.12–17 The characteristics of the RCTs are presented in Table 1. The first 5 studies were included in a 2008 Cochrane Library meta-analysis, with the findings of the planned primary and secondary outcomes presented in Table 2. Although the overall cesarean delivery rate was not significantly different between the groups, the increased rate of diagnosing oligohydramnios, labor induction, and cesarean delivery for fetal distress coupled with the statistically similar rates of primary and secondary outcomes led the authors to conclude that “the SDP measurement seems to be the more appropriate method for assessing AFV during fetal surveillance.” Of note, 2 studies employed lower thresholds for diagnosing oligohydramnios using SDP and higher cutoffs for diagnosing oligohydramnios using AFI.12,13 These important methodological differences must be considered as potential contributors to the increased frequency of diagnosing oligohydramnios and thus intervention when using the AFI compared with the SDP. The authors further noted the need for a systematic review of the diagnostic accuracy of the AFI vs the SDP and additional trials to create outcomes-based consensus in standardizing the method of diagnosing decreased amniotic fluid and timing and mode of delivery in affected pregnancies.18 This
<table>
<thead>
<tr>
<th>Study</th>
<th>Setting</th>
<th>Inclusive years</th>
<th>Subjects (all singleton pregnancies)</th>
<th>Oligohydramnios frequency, n (%)</th>
<th>Primary outcome</th>
<th>Additional outcomes</th>
<th>Fetal testing</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alfirevic et al</td>
<td>United Kingdom</td>
<td>1994—1995</td>
<td>≥41 3/7 wk</td>
<td>SDP AFI</td>
<td>Cesarean delivery</td>
<td>Labor induction</td>
<td>AFV and CTG</td>
<td>Testing every 3 d with labor induction for abnormal CTG or AFV, 43 wk, maternal request, developed exclusion criterion</td>
</tr>
<tr>
<td>Oral et al</td>
<td>Turkey</td>
<td>1997—1998</td>
<td>≥41 3/7 wk</td>
<td>SDP AFI</td>
<td>Not specified</td>
<td>Labor induction</td>
<td>AFV and CTG</td>
<td>Twice weekly testing with delivery for abnormal CTG, oligohydramnios, or 42 wk</td>
</tr>
<tr>
<td>Moses et al</td>
<td>United States</td>
<td>2001—2003</td>
<td>Admitted to labor and delivery and delivery expected (87.5%, ≥37 wk)</td>
<td>SDP AFI</td>
<td>Cesarean delivery for fetal distress</td>
<td>Labor induction</td>
<td>AFI or SDP</td>
<td>Assessment of amniotic fluid by assigned method Results not used for management</td>
</tr>
</tbody>
</table>

Exclusions: Hypertension, Proteinuria, Antepartum hemorrhage, Poor obstetrical history, Ultrasound suspects fetal growth restriction.
### TABLE 1
Characteristics of randomized clinical trials comparing single deepest pocket vs amniotic fluid index (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Setting</th>
<th>Inclusive years</th>
<th>Subjects (all singleton pregnancies)</th>
<th>Subjects assigned to each group (n)</th>
<th>Oligohydramnios definition</th>
<th>Oligohydramnios frequency, n (%)</th>
<th>Primary outcome</th>
<th>Additional outcomes</th>
<th>Fetal testing</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magann et al 15</td>
<td>United States</td>
<td>Not reported</td>
<td>High risk undergoing weekly BPP (43.4%, ≥37 wk)</td>
<td>264 273</td>
<td>&lt;2 cm ≤5.0 cm</td>
<td>46 (17.4) 102 (37.4)</td>
<td>.001</td>
<td>Cesarean delivery for fetal distress</td>
<td>Fetal induction</td>
<td>BPP using AFI or BPP using SDP</td>
</tr>
<tr>
<td>Chauhan et al 16</td>
<td>United States</td>
<td>1997—2001</td>
<td>Medical or obstetrical complications undergoing weekly modified BPP (90.1%, ≥37 wk)</td>
<td>558 530</td>
<td>&lt;2 cm ≤5.0 cm</td>
<td>57 (10.2) 88 (16.6)</td>
<td>.002</td>
<td>Cesarean delivery for nonreassuring fetal heart rate tracing</td>
<td>Nonreassuring fetal heart rate tracing</td>
<td>NST and AFI or NST and SDP</td>
</tr>
</tbody>
</table>

The caveat is reasonable, considering that a total of 500 subjects demonstrated oligohydramnios by either AFI or SDP measurements. More recently, an updated meta-analysis incorporating the results of the latest RCT confirmed the earlier findings of the Cochrane Library report with 1 exception. The updated analysis identified a significantly decreased risk of nonreassuring fetal heart rate tracings in pregnancies evaluated using the SDP (relative risk [RR], 0.85; 95% confidence interval [CI], 0.74–0.97). These authors, like those of the Cochrane meta-analysis, concluded that based on current evidence, MVP should be the method of choice for assessing fluid volume as a measure of fetal well-being. Moreover, they observed that further investigation should assess outcomes in both high-risk and low-risk pregnant women to determine which modality is more effective in preventing poor consequences in either subset population.

The Cochrane Library meta-analysis is the reference cited as supporting professional organizations’ initial expression of preferring SDP vs AFI when diagnosing oligohydramnios. The “Fetal Imaging: Executive Summary of a Joint Eunice Kennedy Shriver National Institutes of Health and Human Development, Society for Maternal-Fetal Medicine, American Institute of Ultrasound in Medicine, American College of Obstetricians and Gynecologists, American College of Radiology, Society of Pediatric Radiology, and Society of Radiologists in Ultrasound Fetal Imaging Workshop” noted that the maximum vertical pocket method for amniotic fluid assessment is preferred because “defining oligohydramnios as a maximum vertical pocket shorter than 2 cm will result in fewer obstetric interventions without a significant difference in the number of patients presenting for labor or preterm evaluation.”

### TABLE 1
Characteristics of randomized clinical trials comparing single deepest pocket vs amniotic fluid index (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Setting</th>
<th>Inclusive years</th>
<th>Subjects (all singleton pregnancies)</th>
<th>Subjects assigned to each group (n)</th>
<th>Oligohydramnios frequency, n (%)</th>
<th>Primary outcome</th>
<th>Additional outcomes</th>
<th>Fetal testing</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kehl et al 17</td>
<td>Germany 2012–2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AFI, amniotic fluid index; AFI, amniotic fluid index; BPP, biophysical profile; CTG, cardiotocogram; NICU, neonatal intensive care unit admission; NST, nonstress test; SDP, single deepest pocket.

perinatal outcomes when compared with an AFI ≤5 cm.8 Subsequent clinical guidance from the American College of Obstetricians and Gynecologists and the American Institute of Ultrasound in Medicine aligns with this stated preference.1 People e3,9,10 The Society for Maternal Fetal Medicine offers a more prescriptive approach in its series “Choosing Wisely - Twenty Things Physicians and Patients Should Question,” stating “Don’t use amniotic fluid index to make a diagnosis of oligohydramnios (in the third trimester).”11 Considering the 2 meta-analyses findings, accompanying qualified conclusions, and authors’ recommendations for additional study, more balanced clinical guidance is warranted.

**Perinatal Outcomes**

The first observation in the rationale for using SDP vs AFI is the absence of significant differences in perinatal outcomes observed with the 2 approaches. The AFV assessments in published RCTs comparing SDP vs AFI were performed in the setting of indicated antepartum fetal surveillance or under research conditions where such surveillance might otherwise occur.12–17 The goal of antepartum fetal surveillance is, quite simply, to prevent stillbirth.3,20 It logically follows that studies comparing SDP vs AFI would select stillbirth or perinatal death as the primary outcome of interest. At first glance, it is surprising that none of the 6 RCTs comparing SDP vs AFI employed either stillbirth or perinatal death as the primary outcome (Table 1). Only 1 study specifically included stillbirth as a secondary outcome, finding no stillbirth in the 530 subjects randomized to AFI and 558 subjects randomized to SDP.16 Of note, 3 RCTs included perinatal mortality as a secondary outcome, also observing no death in either AFI or SDP groups.12,13,17 Another RCT examined neonatal mortality, which was observed in 2 of 273 subjects randomized to AFI assessment and 1 of 264 subjects randomized to SDP.15 The sixth study did not report any measure of fetal or neonatal death.14 Collectively, no stillbirth occurred across 4 RCTs reporting stillbirth, either directly or indirectly via perinatal death, in a total of 1326 subjects assigned to AFI and 1365 subjects assigned to SDP.12,13,16,17 Stillbirths may be categorized as early (20–27 weeks of gestation) or late (≥28 weeks of gestation), with each respective gestational age group accounting for approximately half of stillbirths reported in the United States. We noted that the reported US late fetal death rate is 2.88 per 1000 live births plus fetal deaths at ≥28 weeks of gestation.21 Understanding that this rate is not corrected for lethal congenital anomalies, multiple gestations, and unpredictable unavoidable events, such as placental abruption or umbilical cord prolapse, it nonetheless provides a reasonable frame of reference for the following example. A sample size calculation providing 80% power with an alpha level of .05 to detect a 50% reduction in the late stillbirth rate to 1.44 per 1000 would require randomizing a total of 32,630 subjects to

**TABLE 2**

Results of Cochrane Library meta-analysis examining the AFI vs SDP for preventing adverse pregnancy outcomes18

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Number of studies</th>
<th>Total events</th>
<th>Number of participants</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AFI</td>
<td>SDP</td>
<td>AFI</td>
</tr>
<tr>
<td>Primary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NICU admission</td>
<td>5</td>
<td>157</td>
<td>150</td>
<td>1600</td>
</tr>
<tr>
<td>Perinatal death</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>828</td>
</tr>
<tr>
<td>Secondary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate of diagnosis of oligohydramnios</td>
<td>5</td>
<td>347</td>
<td>153</td>
<td>1600</td>
</tr>
<tr>
<td>Umbilical artery pH&lt;7.1</td>
<td>3</td>
<td>48</td>
<td>44</td>
<td>1302</td>
</tr>
<tr>
<td>5-min Apgar score &lt;7</td>
<td>5</td>
<td>32</td>
<td>28</td>
<td>1600</td>
</tr>
<tr>
<td>Meconium</td>
<td>5</td>
<td>203</td>
<td>191</td>
<td>1600</td>
</tr>
<tr>
<td>Nonreassuring fetal heart rate</td>
<td>4</td>
<td>192</td>
<td>174</td>
<td>1350</td>
</tr>
<tr>
<td>Rate of labor induction</td>
<td>4</td>
<td>154</td>
<td>80</td>
<td>1070</td>
</tr>
<tr>
<td>Assisted vaginal delivery</td>
<td>4</td>
<td>240</td>
<td>228</td>
<td>1552</td>
</tr>
<tr>
<td>Assisted vaginal delivery for distress</td>
<td>2</td>
<td>79</td>
<td>77</td>
<td>803</td>
</tr>
<tr>
<td>Cesarean delivery</td>
<td>5</td>
<td>229</td>
<td>211</td>
<td>1600</td>
</tr>
<tr>
<td>Cesarean delivery for fetal distress</td>
<td>5</td>
<td>100</td>
<td>69</td>
<td>1600</td>
</tr>
</tbody>
</table>

AFI, amniotic fluid index; CI, confidence interval; NE, not estimable; RR, risk ratio; SDP, single deepest pocket.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Number of studies</th>
<th>Total events</th>
<th>Number of participants</th>
<th>Relative effect size</th>
<th>Absolute effect (per 1000)</th>
<th>Grade evidence quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ind</td>
<td>Exp</td>
<td>Ind</td>
<td>Exp</td>
<td>Ind</td>
<td>Exp</td>
</tr>
<tr>
<td>Neonatal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perinatal death</td>
<td>22</td>
<td>4</td>
<td>25</td>
<td>9418</td>
<td>9377</td>
<td>0.31</td>
</tr>
<tr>
<td>Stillbirth</td>
<td>22</td>
<td>2</td>
<td>16</td>
<td>9418</td>
<td>9377</td>
<td>0.30</td>
</tr>
<tr>
<td>NICU admission</td>
<td>17</td>
<td>741</td>
<td>844</td>
<td>8931</td>
<td>8895</td>
<td>0.88</td>
</tr>
<tr>
<td>Neonatal encephalopathy</td>
<td>2</td>
<td>16</td>
<td>23</td>
<td>4440</td>
<td>4411</td>
<td>0.69</td>
</tr>
<tr>
<td>5-min Apgar score &lt; 7</td>
<td>20</td>
<td>87</td>
<td>120</td>
<td>9184</td>
<td>9161</td>
<td>0.73</td>
</tr>
<tr>
<td>Birth trauma</td>
<td>5</td>
<td>40</td>
<td>41</td>
<td>6568</td>
<td>6538</td>
<td>0.97</td>
</tr>
<tr>
<td>Maternal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cesarean delivery</td>
<td>31</td>
<td>1751</td>
<td>1935</td>
<td>10,657</td>
<td>10,373</td>
<td>0.90</td>
</tr>
<tr>
<td>Operative vaginal delivery</td>
<td>22</td>
<td>1331</td>
<td>1244</td>
<td>9436</td>
<td>9148</td>
<td>1.03</td>
</tr>
<tr>
<td>Severe perineal tear</td>
<td>5</td>
<td>189</td>
<td>182</td>
<td>5801</td>
<td>5788</td>
<td>1.04</td>
</tr>
<tr>
<td>Postpartum hemorrhage</td>
<td>9</td>
<td>505</td>
<td>495</td>
<td>6306</td>
<td>6303</td>
<td>1.02</td>
</tr>
<tr>
<td>Breastfeeding at discharge</td>
<td>2</td>
<td>1897</td>
<td>1872</td>
<td>3779</td>
<td>3708</td>
<td>1.00</td>
</tr>
<tr>
<td>Maternal length of stay (d)</td>
<td>7</td>
<td>—</td>
<td>—</td>
<td>2070</td>
<td>2050</td>
<td>—</td>
</tr>
</tbody>
</table>

CI, confidence interval; Exp, expectant management; Ind, induction of labor; RR, risk ratio.

an evidence addressing the issue of perinatal outcomes, including stillbirth and perinatal death, than that examining labor induction, the very intervention cited as favoring using SDP vs AFI.

**Obstetrical Interventions**

The second consideration leading to recommending SDP vs AFI was the increased frequency of diagnosing oligohydramnios and increased labor induction rate associated with using AFI vs SDP. The stated purpose of antepartum fetal surveillance, avoiding stillbirth, raises the expectation that when the test results suggest the risk or presence of fetal jeopardy, further evaluation of the fetal condition or delivery will be indicated. Thus, it is understood that antepartum fetal surveillance is performed only at gestational ages and under circumstances where delivery of a potentially viable newborn is anticipated. A natural and expected consequence of employing antepartum fetal surveillance is that there will be an accompanying increase in the frequency of interventions, such as labor induction or cesarean delivery, as a direct result of responding to abnormal test results. Clinicians should not only expect but welcome this consequence of testing when the increased intervention rate is accompanied by the intended salutary effect of improved perinatal outcomes, specifically the reduction in the stillbirth and perinatal rate supported by population-based evidence.

A Cochrane Library meta-analysis of 34 RCTs, including >21,500 pregnancies, examined the effects of a policy of labor induction at or beyond 37 weeks of gestation compared with a policy of awaiting spontaneous labor indefinitely, until a specific gestational age, or until a maternal or fetal indication for labor induction arose. The primary outcome was perinatal death, defined as the sum of intrapartum fetal deaths and neonatal deaths in the first week of life. The results of the primary maternal and perinatal outcomes are presented in Table 3. A policy of labor induction was associated not only with a clear reduction in perinatal death, stillbirth, neonatal intensive care admission, and 5-minute Apgar score of <7 but also with a reduction in the cesarean delivery rate without affecting the operative vaginal delivery rate. The number of patients needed to treat with labor induction to avoid 1 perinatal death was 544 (95% CI, 441–1042). Moreover, the secondary outcomes favoring labor induction included meconium aspiration syndrome (RR, 0.75; 95% CI, 0.62–0.92) and birthweight of >4000 g (RR, 0.72; 95% CI, 0.54–0.96). Of the 28 outcomes studied, the only ones favoring expectant care were intrapartum analgesia other than epidural or regional anesthesia (RR, 1.11; 95% CI, 1.05–1.18) and length of neonatal stay of ≤2 days (RR, 1.05; 95% CI, 1.02–1.08). There was no observed difference in subgroup analyses of the primary outcomes by gestational age at labor induction (<40, 40–41, >41 weeks of gestation) with regard to either parity or cervix status.

The improved clinical outcomes resulting from a policy of labor induction at ≥37 weeks of gestation should be further considered in the context of associated financial costs and healthcare resource utilization. Of note, 3 trials included in the Cochrane meta-analysis addressed the financial costs associated with labor induction vs expectant care. Although these studies were conducted in different healthcare systems, each reported modestly lower costs with labor induction than expectant management.

A planned secondary analysis of an RCT included in the Cochrane meta-analysis specifically addressed healthcare resource utilization with labor induction vs expectant management. This study randomized low-risk nulliparous women to labor induction at 39 weeks of gestation or expectant management. Women assigned to labor induction underwent considerably fewer antepartum office or hospital visits and less antepartum fetal surveillance testing, laboratory testing, and treatment with a variety of medications and intravenous hydration than women assigned to expectant management. Concerning delivery admission utilization, women assigned to labor induction more often received cervical ripening, oxytocin infusion, and intravenous pressure catheter placement and
experienced longer labors than women managed expectantly. However, the labor induction group was markedly less likely to receive magnesium sulfate infusion and antibiotic infusion and have a post-delivery hospital stay of >2 days or neonatal hospital stay of >2 days than the expectant management group. Following delivery, healthcare resource utilization was similar by study group except for neonates born to expectantly managed women being more likely to have serum bilirubin testing performed. The authors concluded that “these results demonstrate that the health outcome advantages associated with [labor induction] are gained without incurring uniformly greater healthcare resource use.”

In summary, labor induction conferred significant maternal and perinatal benefits at lower cost, with similar healthcare resource utilization, and no discernable adverse consequence compared with expectant management. This body of evidence compellingly argues for, not against, broader support of labor induction at ≥37 weeks of gestation. Therefore, the increased labor induction rate resulting from an amniotic fluid assessment using AFI vs SDP should be interpreted as an opportunity for improving maternal and perinatal outcomes.

Conclusions
Individual studies, meta-analyses, and clinical guidance issued by professional organizations conclude that SDP is preferred to AFI for performing AFV assessments. The basis for this conclusion is 2-fold: (1) the absence of significantly different maternal or perinatal outcomes by assessment method and (2) increased intervention by labor induction following more frequent oligohydramnios diagnoses using AFI. Regarding the first notion, the RCTs forming the basis for preferring the SDP over the AFE are individually and collectively underpowered to support such a firm conclusion. Specifically, sufficient power is noticeably lacking to discern a difference in stillbirth or perinatal death rates, the primary outcomes of interest in pregnancies undergoing antepartum fetal surveillance.

Concerning the second concept, the available data demonstrated significantly improved maternal and perinatal outcomes with a policy of labor induction at or beyond 37 weeks of gestation vs a policy of expectant care. Many of these important outcomes, including reductions in stillbirth and perinatal death, were among those for which RCTs comparing SDP vs AFI found no significant difference. The benefits of labor induction are gained at a nominally lower financial cost and with favorable healthcare resource utilization relative to a policy of expectant care. Therefore, as evidenced by studying sufficient numbers of subjects, an increased labor induction rate, such as that observed with using AFI, should be viewed positively and should support rather than discourage using AFI vs SDP for amniotic fluid assessment. In consideration of the preceding discussion, we offer the following recommendations.

1. Current clinical guidance should be revised, supporting the use of AFI as an acceptable sonographic technique for amniotic fluid assessment, including diagnosing oligohydramnios in singleton pregnancies during the third trimester of pregnancy.
2. The existing Society for Maternal Fetal Medicine Choosing Wisely Item 15 proscribing using the AFI to diagnose oligohydramnios in the third trimester should be withdrawn.
3. Studies to determine outcome-based thresholds of sonographic amniotic fluid measurements for diagnosing oligohydramnios using SDP and AFI are warranted.

REFERENCES
15. Magann EF, Doherty DA, Field K, Chauhan SP, Muffley PE, Morrison JC. Bio-
physical profile with amniotic fluid volume as-
16. Chauhan SP, Doherty DD, Magann EF, Cahanding F, Moreno F, Klausen JH. Amniotic
fluid index vs single deepest pocket technique
during modified biophysical profile: a randomized
clinical trial. Am J Obstet Gynecol 2004;191:
661–7.
deepest vertical pocket or amniotic fluid index as
evaluation test for predicting adverse pregnancy
outcome (SAFE trial): a multicenter, open-label,
randomized controlled trial. Ultrasound Obstet
18. Nabhan AF, Abdelmoula YA. Amniotic fluid
index versus single deepest vertical pocket as a
screening test for preventing adverse pregnancy
outcome. Cochrane Database Syst Rev
2008;2008:CD006593.
Diagnostic utility of maximum vertical pocket
versus amniotic fluid index in assessing amniotic
fluid volume for the prediction of adverse
maternal and fetal outcomes: a systematic re-
view and meta-analysis. J Matern Fetal Neonatal
20. Vintzileos AM. Evidence-based, compared
with reality-based medicine in obstetrics. Obstet
21. Gregory ECW, Drake P, Martin JA. Lack of
change in perinatal mortality in the United States,
22. Freeman RK, Anderson G, Dorchester W.
A prospective multi-institutional study of ante-
partum fetal heart rate monitoring. I. Risk of
perinatal mortality and morbidity according to
antepartum fetal heart rate test results. Am J
24. Goeree R, Hannah M, Hewson S. Cost-
effectiveness of induction of labour versus serial
antenatal monitoring in the Canadian Multicentre
Postterm Pregnancy Trial. CMAJ 1995;152:
1445–50.
induction near term for women aged 35 or over:
an economic evaluation. BJOG 2017;124:
929–34.
Health resource utilization of labor
induction versus expectant management. Am J Obstet Gynecol 2020;222:
369.e1–11.
Labor induction versus expectant management