The Evidence for Abandoning the Amniotic Fluid Index in Favor of the Single Deepest Pocket

Everett F. Magann, C.A.P.T., M.C., U.S.N.,¹ Suneet P. Chauhan, M.D.,² Dorota A. Doherty, Ph.D.,³ Marcia I. Magann, M.S.N., R.N., C.P.N.P.,⁴ and John C. Morrison, M.D.⁴

ABSTRACT

This study assessed whether the amniotic fluid index (AFI) or the single deepest pocket (SDP) is the best technique to estimate amniotic fluid volume. The AFI and SDP were compared to a dye-determined or directly measured amniotic fluid volume. A PUBMED search from 1990 to 2006 was conducted using the search terms “single deepest pocket” or “largest vertical pocket” or “maximum vertical pocket” or “2X1 pocket” AND “amniotic fluid index”. One study compared the AFI and SDP to a dye-determined amniotic fluid volume. There were 1219 publications that used the search term SDP-LVP-MVP versus 4378 using AFI. Twenty publications contained both the AFI and SDP, but only six compared the AFI and SDP. Both the AFI and the SDP poorly identified abnormal amniotic fluid volumes, and neither technique was superior to the other. The AFI identifies a significantly greater number of women as having oligohydramnios versus the SDP but without any difference in perinatal outcomes. Compared with SDP, AFI excessively characterizes a greater number of pregnancies as having oligohydramnios leading to more interventions without improvement in perinatal outcome. The AFI should be abandoned and the SDP used to estimate amniotic fluid volume.

KEYWORDS: Amniotic fluid index (AFI), single deepest pocket (SDP), pregnancy outcome, oligohydramnios

Aberrations of amniotic fluid volume are associated with adverse pregnancy outcomes. Thus it is understandable why the American College of Obstetricians and Gynecologists recommends determination of amniotic fluid volume with every second- and third-trimester ultrasound.¹ The two most commonly and widely used ultrasonic measurements to estimate amniotic fluid volume are the amniotic fluid index (AFI)² and the single deepest pocket technique (SDP).³ These ultrasound measurements of the amniotic fluid volume are helpful in the second trimester, at the time of the “targeted” ultrasound, to identify structural anomalies that may be linked with high or low fluid volumes. These ultrasound measurements are used along with the nonstress test in the modified biophysical profile (BPP, nonstress test, and AFI)⁴ and in the BPP (SDP, fetal movement, fetal tone, fetal breathing, and nonstress test)⁵ to evaluate fetal well-being in pregnancies identified as being at high risk for an adverse outcome.

¹Departments of Obstetrics and Gynecology, Naval Medical Center—Portsmouth, Portsmouth, Virginia; ²Aurora Health Care, West Allis, Wisconsin; ³University of Western Australia, Perth, Australia; ⁴University of Mississippi Medical Center, Jackson, Mississippi.

Address for correspondence and reprint requests: Suneet P. Chauhan, M.D., Aurora Health Care, 8905 W. Lincoln Ave, PAC, West Allis, WI 53227.


Low fluid volumes in pregnancies at risk for intrauteroplacental insufficiency are believed to represent the end stage of a fetus that is now shunting blood to its heart, adrenal gland, and brain at the expense of the rest of the body to preserve life. The consequence of reduced blood flow to the fetal kidneys is a decrease in fetal urine, which is the primary component of amniotic fluid volume in the second half of pregnancy. Dye-dilution techniques and direct measurement of amniotic fluid volume at time of cesarean are invasive, time consuming, cumbersome, and require laboratory support for which the measurement at cesarean delivery only reflects that volume at delivery and cannot be used for serial antenatal testing.

The use of the SDP technique, as part of a formal process to evaluate the amniotic fluid volume of pregnancies at risk for an adverse pregnancy outcome, can be traced back to an article published by Manning in 1980. As a component of the BPP, he described normal amniotic fluid volume as "fluid evident throughout the uterine cavity as well as a largest pocket of fluid measuring more than 1 cm in the vertical dimension." One year later another article published by the same investigators had a new definition of normal amniotic fluid, which was defined as a pocket of fluid that measured at least 1 cm in two perpendicular planes (1 × 1 pocket). By 1990, the definition of an adequate amniotic fluid volume, as a component of the BPP, was a pocket of fluid up to 2 cm in the largest vertical axis with a horizontal component of at least 1 cm (2 × 1 pocket). This change in the fluid component of the BPP was based on a paper by Chamberlain in 1984.7 Chamberlain defined oligohydramnios if the largest vertical pocket was < 1 cm. In all cases the horizontal measurement of the fluid pocket had to be at least 1 cm. He defined a marginal group in which the largest vertical pocket ranged from 1 to 2 cm, and the group from 2 to 8 cm was labeled normal. Corrected perinatal mortality was linked to the measurement of that largest vertical pocket and was 1.97 of 1000 for the 2- to 8-cm group, 37.7 of 1000 for the marginal group of 1 to 2 cm, and 109.4 of 1000 for the group < 1 cm. The investigators believed that the criteria of < 1 cm was too stringent and therefore combined the < 1 cm and the 1- to 2-cm groups together and considered that group to represent pregnancies with oligohydramnios. This threshold of 2 cm is the current ultrasound measurement used to estimate normal amniotic fluid volume, and less than a vertical pocket of 2 cm is used to define oligohydramnios when the BPP is used in antenatal testing.

The AFI was introduced as a four-quadrant evaluation of the amniotic fluid volume in 1987 by Phelan. Rutherford et al. reported that as the AFI decreased, the number of nonreactive nonstress tests increased, and this finding was significant at an AFI of ≤ 5. She also observed that the risk of meconium-stained amniotic fluid, cesarean delivery for fetal intolerance of labor, and 1- and 5-minute Apgar scores < 6 were significantly more frequent in the women with an AFI ≤ 5 compared with women with an AFI > 5. The BPP using the SDP as the amniotic fluid volume evaluation and the nonstress test with the AFI used to estimate amniotic fluid volume were both used to evaluate at-risk pregnancies until an article in 1990 that claimed the superiority of the AFI versus the SDP technique in the estimation of amniotic fluid volume. Moore evaluated 791 women with both the AFI and SDP techniques. The critical values (5th and 95th percentiles) had been calculated for both the AFI and SDP technique. The AFI was used as the volume standard based on an earlier study in six sheep (Moore TR, Brace RA. Amniotic fluid index in the term ovine pregnancy: A predictable relationship between AFI and amniotic fluid volume [Abstract No. 286]. Meeting of the Society for Gynecological investigation, Baltimore MD, March 17–20, 1988). In that investigation, all of the amniotic and allantoic fluid was drained from six sheep and then saline was infused in 100-mL increments and AFI measurements were taken. Curve-fitting formulas were used and a close linear relationship was observed between the AFI and amniotic fluid volume. In his evaluation of 791 pregnant women, Moore identified 76 women with oligohydramnios using the AFI measurement and 32 with oligohydramnios with the SDP technique. He concluded that because the AFI had identified more pregnancies with oligohydramnios than the SDP technique, it was a better predictor of oligohydramnios.

This article had a major impact on how amniotic fluid volume is estimated. In a PUBMED search from 1990 to 2006 using the search terms “single deepest pocket” or “SDP” or “largest vertical pocket” or “LVP” or “maximum vertical pocket” or “MVP” or “2 × 1 pocket,” 1219 articles were identified. Using the search terms “AFI” or “amniotic fluid index,” 4378 articles were identified.

If either of the two techniques is to be acknowledged superior to the other, it must meet most, if not all of the following criteria: (1) Accurately identify the actual amniotic fluid (analysis by dye dilution technique or direct measurement at cesarean), (2) Among uncomplicated pregnancies with reassuring outcomes, minimize the number of patients it considers to have oligohydramnios or hydramnios, (3) Is a reliable predictor of adverse peripartum complications, and (4) When used in conjunction with other ancillary tests, it improves the neonatal outcomes.

ASSESSMENT OF AMNIOTIC FLUID VOLUME WITH AFI VERSUS SDP

The comparisons of the AFI and the SDP are problematic because this is a comparison of two sonographic
measurements without any objective criteria or gold standard assessment. If the comparison is truly to determine if the AFI is superior to the SDP, then both must be compared with a dye-determined technique or direct measurement of amniotic fluid volume. That assessment was undertaken and published in 2000. Magann et al estimated amniotic fluid volume by AFI and the SDP techniques among 179 patients prior to an ultrasound directed amniocentesis with dye-dilution and spectrophotometric calculation of the actual amniotic fluid volume. Using sensitivities and specificity positive and negative predictive values and likelihood ratios, no difference was noted in the AFI and the SDP to identify abnormal amniotic fluid volumes. Both techniques identified abnormal fluid volumes so poorly that they were determined to be unreliable. Others investigators have tried to correlate the AFI and the SDP measurements with dye determination (at the time of amniocentesis) or directly measured amniotic fluid volume (at the time of cesarean delivery) and have observed the poor predictive value of the AFI and SDP to predict abnormal amniotic fluid volumes (oligohydramnios or polyhydramnios).

Magnetic resonance imaging (MRI) was compared with the AFI and the SDP and then correlated with directly measured amniotic fluid volume at the time of delivery. The MRI was superior to the SDP and equivalent to the AFI as determined by Pearson correlation. However, when the ability to predict oligohydramnios was assessed, MRI, AFI, and SDP techniques were all similar.

**Normative Data of AFI vs SDP Among Uncomplicated Pregnancies**

In 1990, Moore and Cayle provided the first gestational age–based normative data for AFI among pregnancies, with no known complications, no ultrasonographic abnormalities, term delivery, birthweight between the 10th and 90th percentile, and Apgar score > 6 at 5 minutes. They concluded that AFI values for each week were distinct, and week-specific tables were needed. Their conclusion has been shown to be based on an uninterruptible equation, and their cutoffs to label a pregnancy as having an abnormal amniotic fluid volume is not reproducible in different patient populations.

In 1993, Nwosu and colleagues undertook an assessment of the AFI across gestation using longitudinal evaluations rather than the cross-sectional method used by Moore. They evaluated the AFI in 105 pregnancies. Each pregnancy was evaluated six times at 4-week intervals across gestation. They observed marked differences at the normal thresholds for the lower limits of the AFI. Using the normal ranges by Nwosu compared with the Moore data, significantly fewer women would be labeled with oligohydramnios.

In trying to confirm the normative data published by Moore and Cayle and the data published by Nwosu et al, we assessed AFI and SDP among 1400 singleton pregnancies with identical inclusion criteria. Because all pregnancies were uncomplicated and had a normal outcome, the sonographic technique that classified the least number of patients as having abnormal amniotic fluid should be the preferred technique. Using AFI ≤ 5.0 cm as the criteria to diagnose oligohydramnios, 8% of the pregnancies would be classified as having low amniotic fluid volume; with SDP < 2.0 cm, 1% would be labeled with low fluid volumes in this normal patient population. For hydramnios (AFI ≥ 24 cm, SDP ≥ 8.0 cm), the corresponding values were 0% and 0.7%, respectively. Gramellini et al undertook a prospective cross-sectional study to construct normal reference ranges of four ultrasound parameters for the evaluation of amniotic fluid volume in the second trimester of pregnancy. Using ultrasound, he assessed the estimated amniotic fluid volumes between 11 and 24 weeks of gestation in 273 normal pregnant patients. The largest amniotic pocket in a vertical dimension free of small parts and umbilical cord was measured. The maximum vertical and transverse diameters on the same scan were measured along with the mean diameter, and the product of the two diameters was calculated. All four of these measurements were analyzed and had good intra- and interoperator reproducibility showing significant correlation with gestational age and biparietal diameter that increased between 12 and 24 weeks. They concluded that a semiquantitative method based on a single amniotic fluid pocket correlated with gestational age and may be a more appropriate method to determine amniotic fluid volume abnormalities prior to 23 to 24 weeks of gestation.

**AFI VERSUS SDP AS PREDICTOR OF ADVERSE OUTCOMES**

A small number of investigations have compared the AFI and the SDP technique. A PUBMED search using the search terms “amniotic fluid index,” “single deepest pocket,” “maximum vertical pocket” discovered 20 publications but only 12 publications that have used both techniques to predict peripartum complications (Tables 1 and 2). Some of these reports use different definitions of oligohydramnios, and they also assessed varying outcomes of interest. For example, the study by Miyamura used a SDP ≤ 3 and an AFI ≤ 8 to define oligohydramnios. Myles evaluated not only the AFI and SDP but also the amniotic fluid volume distribution and used ≤ 2.5 cm rather than ≤ 2 cm for defining oligohydramnios using the SDP technique.
Ten years after the initial publication by Phelan et al describing AFI, Miyamura et al\textsuperscript{22} compared the AFI with the SDP in a Japanese population. They compared the predictability of the last measurement of the AFI and SDP, which was done on all of the 69 singleton pregnancies that were undergoing antenatal testing and who were small for gestational age. They found that the SDP and AFI were significantly correlated but that a SDP of < 3 cm was the most useful criteria for the presence of oligohydramnios that may result in fetal distress in labor. Seven subsequent nonrandomized trials have reached varying conclusions when comparing the AFI versus SDP to predict adverse pregnancy outcomes. The AFI has been reported as being the superior test by some investigators\textsuperscript{23,24} the SDP the superior test by other investigators\textsuperscript{27} and neither test was superior to the other or accurately identifies perinatal complications in

<table>
<thead>
<tr>
<th>N</th>
<th>Peripartum Complications of Interest</th>
<th>Oligo</th>
<th>Results</th>
<th>Recommend</th>
</tr>
</thead>
<tbody>
<tr>
<td>69</td>
<td>Fetal labor intolerance in SGA infants</td>
<td>SDP—13 (19%)</td>
<td>SDP &lt; 3 is better predictor of intrapartum fetal distress than AFI &lt; 8</td>
<td>SDP</td>
</tr>
<tr>
<td>266</td>
<td>Fetal labor intolerance, admission to NICU</td>
<td>SDP—38 (14%)</td>
<td>AFI is better predictor of CD for fetal distress than SDP compared with SDP &lt; 1 cm, AFI ≤ 5.0 cm has greater sensitivity but similar specificity, positive and negative predictive values for IUGR, fetal intolerance of labor, distress during labor, low Apgar score at 5 minute, and perinatal mortality</td>
<td>AFI*</td>
</tr>
<tr>
<td>174</td>
<td>Fetal intolerance of labor, low Apgar score, IUGR, perinatal mortality</td>
<td>SDP—49 (40.2%)</td>
<td>AFI—70 (28.2%)</td>
<td>AFI</td>
</tr>
<tr>
<td>100</td>
<td>UA pH &lt; 7.20</td>
<td>SDP—0 (0%)</td>
<td>Correlated AFI and SDP with dye-determined amniotic fluid volumes. Both techniques poorly identified low fluid volume: AFI 3/24 and SDP 0/24. Neither AFI nor SDP identified neonatal acidosis</td>
<td>Neither</td>
</tr>
<tr>
<td>291</td>
<td>Fetal labor intolerance, low Apgar score, UV pH &lt; 7.15, admission to NICU</td>
<td>SDP—14 (4.8%)</td>
<td>Using ROC curves, both techniques are of limited value and neither AFI nor SDP accurately identified pregnancies at risk for an adverse outcome</td>
<td>Neither</td>
</tr>
<tr>
<td>198</td>
<td></td>
<td>SDP—11 (5.5%)</td>
<td>Using ROC curves, SDP of 2.7 cm was better predictor of perinatal complications than AFI</td>
<td>SDP</td>
</tr>
<tr>
<td>1584</td>
<td>Fetal intolerance of labor, meconium aspiration, birth asphyxia, NICU admission</td>
<td>SDP—22 (1.4%)</td>
<td>Although better than SDP, AFI has poor sensitivity for predicting perinatal complications</td>
<td>Neither</td>
</tr>
<tr>
<td>41</td>
<td>Fetal intolerance of labor</td>
<td>SDP—7 (17%)</td>
<td>Neither an AFI ≤ 5 or a SDP ≤ 2 are better than the normal measurement AFI &gt; 5 or SDP &gt; 2 in the prediction of fetal intolerance of labor</td>
<td>Neither</td>
</tr>
</tbody>
</table>

* The AFI and deepest vertical pocket had similar sensitivity for caesarean section (23% and 15%, respectively) and positive predictive value (8% and 5%).

Oligo, oligohydramnios; SGA, small for gestational age; SDP, single deepest pocket; AFI, amniotic fluid index; NICU, neonatal intensive care unit; CD, caesarean; IUGR, intrauterine growth restriction; UA, umbilical arterial; UV, umbilical vein; ROC, receiver operating characteristic.
other investigations (Table 1). The trials evaluated estimated amniotic fluid volumes immediately prior to delivery and correlated that volume with intrapartum pregnancy outcomes. One trial estimated amniotic fluid volumes to intrapartum pregnancy outcomes within 7 days. The final four trials evaluated amniotic fluid volumes as a component of an antenatal test (BPP or modified BPP) and linked these with pregnancy outcomes. Except for the Morris trial, the major shortcomings of these trials was their small size (< 300). There were different definitions of neonatal acidosis among the trials. The Morris trial had a calculated sample size of 4500 participants but was only able to enroll 1584 women because of change in the practice management of women with prolonged pregnancies. Although this trial concluded that the AFI was superior to the SDP in evaluating the amniotic fluid volumes after 40 weeks, the sensitivity of the AFI was poor for an adverse pregnancy outcome. The authors came to the conclusion that the routine use of the AFI in pregnancies beyond 40 weeks is likely to lead to increased interventions without improvement in pregnancy outcomes.

We found four randomized clinical trials (RCTs) that have compared the accuracy of identifying peripartum complications using AFI versus SDP (Table 2). Three of the trials evaluated the amniotic fluid volumes during antenatal testing and one upon admission to labor and delivery (intrapartum assessment). All four of these RCT reached the same conclusion: Compared with SDP, AFI is significantly more likely to categorize patients as having oligohydramnios, without a concomitant improvement in the perinatal outcomes. Admittedly, these four studies did not have sufficient sample size to detect a meaningful difference in peripartum complications. The fact that use of SDP leads to significantly less diagnosis of oligohydramnios and is not linked to increased morbidity is sufficient a reason to adopt it as the method to assess amniotic fluid.

### Table 2 Randomized Clinical Trials Comparing Single Deepest Pocket (SDP) and Amniotic Fluid Index (AFI)

<table>
<thead>
<tr>
<th>Population</th>
<th>SDP</th>
<th>AFI</th>
<th>Results</th>
<th>Recommend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alfirevic et al</td>
<td>SDP—2% AFI—10% (p = 0.0008)</td>
<td>Significantly more induction in the group randomized to AFI, without improved perinatal outcomes</td>
<td>SDP</td>
<td></td>
</tr>
<tr>
<td>Magann et al</td>
<td>SDP—17% AFI—38% (p &lt; 0.0001)</td>
<td>Significantly more induction in the group randomized to AFI, without improved perinatal outcomes</td>
<td>SDP</td>
<td></td>
</tr>
<tr>
<td>Chauhan et al</td>
<td>SDP—10% AFI—17% (p = 0.002)</td>
<td>No difference in the peripartum outcomes</td>
<td>SDP</td>
<td></td>
</tr>
<tr>
<td>Moses et al</td>
<td>SDP—8% AFI—25% (p = 0.001)</td>
<td>No difference in the peripartum outcomes</td>
<td>SDP</td>
<td></td>
</tr>
</tbody>
</table>

SDP, single deepest pocket; AFI, amniotic fluid index; Oligo, oligohydramnios; GA, gestational age.

### AFI VERSUS SDP: USED IN CONJUNCTION WITH OTHER ASSESSMENT OF FETAL WELL-BEING

As part of the modified BPP, the AFI has been used with the nonstress test to assess fetal well-being. In a randomized trial, the modified BPP was compared with the contraction stress test (CST) to identify those patients at increased risk for adverse perinatal outcomes in pregnancies complicated with small-for-gestational-age fetuses. There was no significant benefit observed with the CST was compared with the biophysical profile as a backup test, and the CST was associated with a higher rate of intervention for an abnormal test than was the biophysical profile. Other investigators have reported that the false-positive rate with the modified BPP is similar to CST, and the rate of iatrogenic prematurity from intervention for false-positive test results occurred in 1.5% of women tested being delivered prior to 37 weeks of gestation. The shortcomings with modified BPP are as follows: (1) They have not been linked to umbilical arterial acid-base abnormalities (which are an objective assessment of fetal well-being), (2) outcomes have been assessed in only a small number of cohorts, and (3) they have not been correlated to the prevention of long-term morbidity, such as cerebral palsy.

Publications on the complete BPP have addressed the shortcomings cited for the modified BPP. Manning et al obtained BPP prior to cordocentesis among 493 cohorts with intrauterine growth restriction or alloimmune anemia. He noted a highly significant linear correlation between BPP score and umbilical venous pH. Poor BPP score performance (a score of 0 of 10) was always associated with a pH < 7.20, whereas the pH was always > 7.20 when the biophysical profile score was 10 of 10. Although the reports using the modified biophysical profile have experience with < 16,000 cohorts, Dayal et al have reported on the experience with > 80,000 patients managed with BPP > 18 years and noted that fetal deaths after a reassuring BPP are
primarily due to events that occurred after the antepartum testing. Finally, using the SDP, Manning et al.\textsuperscript{13} reported that the rate of cerebral palsy is significantly lower among patients managed with BPP than those without it. Admittedly with the use of the BPP in antenatal testing, it is not the assessment of amniotic fluid with SDP alone that is responsible for the improved outcomes but the combination of the five interrelated components. However, if one was to select two components (assessment of amniotic fluid and the fetal heart rate tracing) of the BPP and use them as a stand-alone test, we should not modify them. Because the criteria of the reactive nonstress test are similar when used as part of modified and complete BPP, the method to assess amniotic fluid should be the same, namely the SDP technique.

In conclusion, we cannot find any objective reason to favor AFI over SDP. It appears that the AFI compared with the SDP excessively characterizes patients as having oligohydramnios leading to an increase in obstetric interventions without any documented improvement in perinatal outcome. Appropriately designed and powered studies are needed to determine the best method to evaluate the amniotic fluid volume. Until there is evidence of the objective superiority of the four-quadrant summation of amniotic fluid compared with the SDP technique, we recommend abandoning the AFI in favor of the SDP.

DISCLOSURE

The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, or the United States Government. We are military service members. This work was prepared as part of our official duties. Title 17 U.S.C. 105 provides that “Copyright protection under this title is not available for any work of the United States Government.” Title 17 U.S.C. 101 defines a United States Government work as a work prepared by a military service member as part of that person’s official duties.

REFERENCES