Interexaminer Agreement in Physical Examination for Children With Suspected Soft Tissue Abscesses

Joan Elizabeth Giovanni, MD, Mary Denise Dowd, MD, MPH, Chris Kennedy, MD, and Jeffrey G. Michael, DO

**Objective:** This study aimed to measure interexaminer agreement for physical examination (PE) findings in children with a suspected soft tissue abscess.

**Methods:** A prospective study was conducted from March 1 to July 31, 2007, at an urban, tertiary care children's hospital emergency department. Children presenting to the emergency department with a suspected local skin abscess were independently examined by 2 physicians. Interrater agreement of 7 PE findings for children with a suspected soft tissue abscess was assessed. Interrater agreement was calculated for the diagnosis of the lesion and decision to incise and drain.

**Results:** A total of 105 paired observations were completed by a total of 27 physicians. The patients examined were aged 2 weeks to 18 years, with a mean age of 80 months. Lesions were most frequently encountered on the buttocks (38%). Incision and drainage was attempted in 75% of cases, with purulent material obtained in 92% of all attempts. Interrater agreement was substantial for erythema ($\kappa = 0.66$) and size of the lesion (intraclass correlation coefficient = 0.78), moderate for drainage ($\kappa = 0.57$) and tenderness ($\kappa = 0.40$), fair for fluctuance ($\kappa = 0.35$), and poor for warmth ($\kappa = 0.15$) and showed no agreement for induration ($\kappa = -0.08$). There was moderate agreement on diagnosing the lesion as an abscess ($\kappa = 0.48$) and determination if the lesion required incision and drainage ($\kappa = 0.44$).

**Conclusions:** Interexaminer agreement of examination findings and diagnosis of an abscess was fair to moderate, implying a lack of precision of PE as the primary means for diagnosis. Future studies of diagnostic adjuncts, such as bedside ultrasonography, may lead to improved management of soft tissue infections in children.

**Key Words:** abscess, interexaminer agreement, physical examination, $\kappa$

(Pediatr Emer Care 2011;27: 475–478)

Pediatric emergency department (ED) visits for skin and soft tissue infections have increased dramatically during the past decade, largely attributable to the increase of methicillin-resistant *Staphylococcus aureus* in the community. An abscess is defined as a localized collection of pus caused by infection, whereas cellulitis is considered an infection of the skin and subcutaneous tissue. Distinguishing the two on physical examination in children can be difficult because findings overlap. The clinical management, however, differs between cellulitis and an abscess because cellulitis is typically treated with antibiotics and an abscess is usually treated with incision and drainage (I&D).

Currently, the most common diagnostic practice used to distinguish an abscess from cellulitis is physical examination, with certain findings such as fluctuance considered more consistent with abscess. However, no studies have examined interexaminer reliability of physical examination for the diagnosis of cellulitis or abscess, especially in children. Interexaminer reliability is dependent on the ability of 2 or more individuals to be consistent. Reliable diagnosis of pediatric soft tissue infections is necessary to provide appropriate patient care and ultimately not subject children to unnecessary procedures. The aim of our study was to assess the interexaminer agreement among physicians for physical examination findings in children with a suspected soft tissue abscess.

**METHODS**

**Study Design**

This cross-sectional study measured physicians’ interrater agreement of physical examination findings for children presenting to the ED with a local skin infection suspected to be an abscess. This study was approved by the hospital’s institutional review board with a waiver of consent because care was not altered and no patient identifiers were collected.

**Study Setting and Population**

The study took place during a 4-month (March 1 to July 31, 2007) period at an urban, tertiary care children’s hospital ED with an annual census of 65,000. Children aged 0 to 18 years with a suspected local skin abscess, as identified by the treating physician, were eligible for observation. Exclusions included those children (1) with previous attempts to incise and drain the presenting lesion, (2) presenting for follow-up care of an abscess, and (3) with perineal infections owing to the possible need for sedation during examination or consultation with a surgical subspecialist. A convenience sample of observations was collected. This method was used because the study required 2 physicians to complete each data collection set—our ED has single-physician coverage for 7 hours each day. Thus, observations were made on eligible patients 7 days a week covering 17 hours each day during the study period. Physicians participating as examiners in the study were full-time or part-time pediatric emergency medicine attending physicians or pediatric emergency medicine fellows. Resident physicians and nurse practitioners were not included in the study.

**Study Protocol**

A standardized data collection tool (available on request) was developed by the primary investigator. The tool consisted of paired collection forms, one for each rater, to document size of the lesion, description of the lesion, and diagnosis and treatment decisions. Two physicians (the treating physician as well as an additional physician working in the ED) completed each paired data collection form. In addition, the treating physician collected additional data including patient age, location of the...
infection, duration of symptoms, whether I&D was performed, and, if performed, whether purulent material was obtained. The data collection forms were reviewed for ease of use by ED physicians before finalization.

Children were initially evaluated by the treating physician according to their usual practice. The first physician, the treating physician, verified inclusion/exclusion criteria and completed a data collection form. Before any medical intervention or other diagnostic tests, a second physician, blinded to the first's assessment, independently evaluated the same child and recorded findings on an identical data collection form within 30 minutes of the treating physician's examination. The second physician was not restricted from obtaining pertinent history during this examination.

**Measurements**

Each physician recorded his/her coded identifier and measured the largest diameter of the lesion to the nearest 0.5 cm with a provided flexible ruler. Identical rulers were used by all participants. Each physician answered “yes” or “No” to the presence or absence of each of the following: erythema, warmth, tenderness, drainage, induration, and fluctuance. The physicians then provided a “yes” or “no” response to the following questions: “Is this lesion an abscess?” and “Would you incise and drain this lesion?”

Physicians were instructed not to share their findings or discuss the plan of patient care before completion of the documentation forms. The physicians independently returned the completed data collection forms to a designated location in the ED. The data collection forms were numbered to collate corresponding sets before data analysis. Participation by the physicians in the study was voluntary.

**Data Analysis**

A sample size of 100 completed data collection sets was determined by examining the expected width of the 95% confidence intervals (CIs) for κ based on normal distribution and asymptotic variance. This sample size would provide confidence limits of ±0.18 for moderately skewed responses and ±0.29 for extremely skewed responses. Such CIs are considered good precision.5 Statistical analyses were designed in consultation with a hospital biostatistician.

Pairwise comparisons were evaluated, and percent for each finding was calculated. To express interrater agreement, the κ statistic was used for dichotomous variables, and the intraclass correlation coefficient (ICC) was used for the size of the lesion, measured as a continuous variable.

A κ or an ICC with a value of 0.8 or greater is considered an almost perfect chance-adjusted agreement, whereas a value of less than 0.2 is a poor chance-adjusted agreement. Substantial chance-adjusted agreement values fall between 0.6 and 0.8, and moderate chance-adjusted agreement values are between 0.4 and 0.6. Values less than zero are considered to have no agreement.6 Data entry and analysis were done with SPSS for Windows (SPSS, Inc, Chicago, Ill) and SAS (SAS Institute, Inc, Cary, NC) software.

**RESULTS**

One hundred five paired observation sets were collected during the 4-month study period. Twenty-seven physicians participated in the study. Patients had a mean (SD) age of 80.6 (65.4) months (range, 0.5–224 months). The presenting lesion, by parental report, was present for a mean (SD) of 3.8 (2.9) days (range, 1–17 days). Most commonly, the lesions evaluated were found on the buttocks (38%), followed by arm/axilla (14%), and thigh (14%).

Of the 105 lesions observed, the treating physicians performed an I&D in 75% (79/105) of cases. Purulent material was obtained in 92% of the procedures (73/79). Subanalysis of κ for various provider combinations (attending-attending, attending-fellow, fellow-fellow) demonstrated similar κ values between the 3 groups (data not shown). However, these smaller sample sizes

---

**TABLE 1. Interrater Chance-Adjusted Agreement for Physical Examination Findings**

<table>
<thead>
<tr>
<th>Interrater Agreement Variable</th>
<th>Value of κ/ICC*</th>
<th>95% CI of κ/ICC*</th>
<th>Absolute Agreement, %</th>
<th>95% CI of Absolute Agreement, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of lesion</td>
<td>0.78*</td>
<td>0.69 to 0.84*</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Erythema</td>
<td>0.66</td>
<td>0.42 to 0.89</td>
<td>93</td>
<td>89–98</td>
</tr>
<tr>
<td>Drainage</td>
<td>0.57</td>
<td>0.40 to 0.74</td>
<td>81</td>
<td>73–89</td>
</tr>
<tr>
<td>Tenderness</td>
<td>0.40</td>
<td>0.06 to 0.73</td>
<td>92</td>
<td>87–98</td>
</tr>
<tr>
<td>Fluctuance</td>
<td>0.35</td>
<td>0.16 to 0.54</td>
<td>71</td>
<td>63–80</td>
</tr>
<tr>
<td>Warmth</td>
<td>0.15</td>
<td>−0.08 to 0.38</td>
<td>79</td>
<td>71–87</td>
</tr>
<tr>
<td>Induration</td>
<td>−0.08</td>
<td>−0.12 to −0.04</td>
<td>85</td>
<td>78–92</td>
</tr>
<tr>
<td>Is lesion an abscess?</td>
<td>0.48</td>
<td>0.19 to 0.77</td>
<td>91</td>
<td>86–97</td>
</tr>
<tr>
<td>Does lesion require I&amp;D?</td>
<td>0.44</td>
<td>0.23 to 0.66</td>
<td>82</td>
<td>75–89</td>
</tr>
</tbody>
</table>

NA indicates not applicable
in the subanalysis were not powered accordingly and preclude any meaningful interpretations without the possibility of introducing a type 2 error.

**DISCUSSION**

Accurate diagnosis of skin and soft tissue infections is a clinical challenge owing to the overlapping physical characteristics of these infections. Common physical examination findings of both cellulitis and an abscess include warmth, erythema, and tenderness of the involved area. Induration is suggestive of an active inflammatory process but not specific for the presence of an abscess. Fluctuance and drainage from a wound are highly suggestive of an abscess but are not always present on physical examination. Our study found that many aspects of the physical examination lack reliability (as defined by interexaminer agreement) and that examiners do not infrequently disagree on whether a lesion is an abscess and whether an I&D needs to be performed.

Concordance for presence of absence of these physical examination findings in children has not been previously reported. For adults presenting to the ED with cellulitis, a previous study examined interrater agreement for examination characteristics of cellulitis. Agreement was strong for objective measurements such as fever or diameter of erythema but poor or moderate for subjective impressions of impression of severity. Similarly, the results from this study demonstrate that chance-adjusted agreement was stronger for more objective measurements, such as lesion size, and weaker for more subjective findings, such as tenderness or induration.

Incision and drainage is the preferred treatment of the management of a skin abscess. An I&D procedure was performed on 75% of our patients, and purulent material was evacuated in 92% of all attempts, suggesting that most of the lesions observed in this study were abscesses. There was no association between the treating physician and the second physician regarding the decision to perform an I&D, the presence of purulent material if an I&D was performed, or any other treatment decision as the physicians independently completed their observations and did not discuss treatment decisions about each patient.

Owing to the painful nature of the procedure, I&D of an abscess in children usually requires the use of sedation. Sedation can be time-consuming for the physician and the family and, most importantly, is not without risk to the child. Thus, accurate diagnosis of an abscess before any invasive procedure or administration of sedative medications is critically important. Given that there is poor or fair chance-adjusted agreement among physicians on the physical examination findings on which the decision to do the procedure is made (fluctuance, induration, warmth), other diagnostic modalities may need to be considered.

Bedside ultrasonography in the ED can be a useful adjunct in the diagnosis and management of skin and soft tissue infections. In one study, bedside ultrasonography changed the ED physician management of cellulitis in nearly half of all cases studied. However, ultrasound equipment is expensive, not available in every ED, and requires specific training and experience.

This study emphasizes that continued education and training may be needed for the diagnosis of a soft tissue abscess in children. For those institutions without the training or availability to perform bedside ultrasonography, consideration should be made to improve physician’s physical examination skills of pediatric soft tissue infections in the ED. Simulated abscess models have been developed to teach I&D procedures. Such models could also be used to aid in physical examination and diagnosis of an abscess not only for students and residents but also for fellows and attending physicians as well.

The small sample sizes in the subanalysis by experience of provider (attending vs fellow) prevented the ability to comment on whether years of experience influenced physical examination skills and agreement, although this small subanalysis failed to show a significant difference between the groups. Future studies may be needed to better clarify if providers with more pediatric ED experience have improved interrater agreement than trainees, such as residents and fellows.

The χ statistic measures agreement beyond that expected by chance alone. Absolute agreement for the variables was quite substantial, with agreement ranging from 71% to 93%. However, chance-adjusted agreement was much poorer and did not reflect the absolute agreement, with wide ranging values from −0.08 to 0.78. Because of these conflicting results, one must consider that the poor χ values may reflect very high levels of expected agreement and skewed marginal frequencies.

There are several limitations to this study. The treating physician, as opposed to the second rater, might have obtained a more detailed history regarding the skin lesion being examined. This potential bias could have been reduced by using 2 physicians not involved in the care of the patient. This was not feasible given limited physician resources at our institution. This was a single-center study and may not be applicable to other settings or institutions. In addition, the assessors could have improved their physical examination skills of abscesses during the period of the study. This bias is likely of little significance because the study was conducted during a short period and in an ED that frequently treats skin infections in children. The use of the χ statistic also has its limitations. χ is a conservative measure and may be low despite high levels of agreement and accurate individual ratings.

**CONCLUSIONS**

The decision to perform an I&D is frequently based on the presence of findings suggestive of an abscess (eg, fluctuance, warmth, erythema). However, fluctuance, arguably the most important clinical finding that distinguishes an abscess from cellulitis, had only fair chance-adjusted agreement and lower absolute agreement. Interrater agreement was moderate for the clinical diagnosis of an abscess. Agreement for physical examination findings in soft tissue infections was substantial for only 2 clinical variables—presence or absence of erythema and lesion size. Our results suggest that physical examination alone may not be sufficient for the accurate diagnosis of an abscess among pediatric ED physicians. Diagnostic adjuncts, such as bedside ultrasonography, may be useful to improve diagnostic accuracy and management of soft tissue infections in children.

**REFERENCES**


