

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

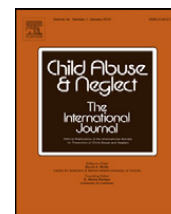
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at SciVerse ScienceDirect

Child Abuse & Neglect



Diagnostic accuracy in child sexual abuse medical evaluation: Role of experience, training, and expert case review

Joyce A. Adams^{a,*}, Suzanne P. Starling^b, Lori D. Frasier^c, Vincent J. Palusci^d,
Robert Allan Shapiro^e, Martin A. Finkel^f, Ann S. Botash^g

^a University of California, San Diego School of Medicine, Rady Children's Hospital, San Diego, CA, USA

^b Eastern Virginia Medical School, Children's Hospital of the King's Daughters, Norfolk, VA, USA

^c University of Utah, Division of Child Protection and Family Health, Salt Lake City, UT, USA

^d NYU School of Medicine, Bellevue Hospital, New York, NY, USA

^e University of Cincinnati College of Medicine, Center for Safe and Healthy Children, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

^f University of Medicine and Dentistry, New Jersey, School of Osteopathic Medicine, Stratford, NJ, USA

^g SUNY Upstate Medical University, Syracuse, NY, USA

ARTICLE INFO

Article history:

Received 23 April 2011

Received in revised form

26 December 2011

Accepted 11 January 2012

Available online 25 May 2012

Keywords:

Child sexual abuse

Medical examination

Interpretation of findings

Competence

ABSTRACT

Objectives: (1) The purpose of this study was to assess the ability of clinicians who examine children for suspected sexual abuse to recognize and interpret normal and abnormal ano-genital findings in magnified photographs using an online survey format. (2) Determine which factors in education, clinical practice, and case review correlate with correct responses to the survey questions.

Methods: Between July and December 2007, medical professionals participated in a web-based survey. Participants answered questions regarding their professional background, education, clinical experience, and participation in case review. After viewing photographs and clinical information from 20 cases, participants answered 41 questions regarding diagnosis and medical knowledge. Answers chosen by an expert panel were used as the correct answers for the survey.

Results: The mean number of correct answers among the 141 first-time survey respondents was 31.6 (SD 5.9, range 15–41). Child Abuse Pediatricians (CAP) had mean total scores which were significantly higher than Pediatricians (Ped) (34.8 vs. 30.1, $p < 0.05$) and Sexual Assault Nurse Examiners (SANE) (34.8 vs. 29.3, $p < 0.05$). The mean total scores for Ped, SANE, and Advanced Practice Nurses (APN) who examine fewer than 5 children monthly for possible CSA were all below 30. Total score was directly correlated with the number of examinations performed monthly ($p = 0.003$). In multivariable regression analysis, higher total score was associated with self-identification as a CAP, reading *The Quarterly Update* newsletter ($p < 0.0001$), and with quarterly or more frequent expert case reviews using photo-documentation ($p = 0.0008$).

Conclusions: Child Abuse Pediatricians, examiners who perform many CSA examinations on a regular basis, examiners who regularly review cases with an expert, and examiners who keep up to date with current research have higher total scores in this survey, suggesting greater knowledge and competence in interpreting medical and laboratory findings in children with CSA. Review of cases with an expert in CSA medical evaluation and staying up to date with the CSA literature are encouraged for non-specialist clinicians who examine fewer than 5 children monthly for suspected sexual abuse.

© 2012 Elsevier Ltd. All rights reserved.

* Corresponding author address: UCSD Pediatric Associates, 7910 Frost Street, Suite 350, San Diego, CA 92123 USA.

Introduction and background

The medical evaluation of children for concerns of sexual abuse requires specialized skills and training. This includes forensic interview skills, knowledge of developmental anatomy and variations in normal anatomy, training to detect conditions that mimic trauma or sexually transmitted infection, and experience in identifying accidental and non-accidental trauma (Kellogg & Committee on Child Abuse and Neglect, 2005). The examination techniques are specialized and unique to the field and are currently performed by physicians, nurse practitioners, physician assistants, and nurses who have varying levels of training, experience, and supervision.

Surveys have shown that practicing physicians and residents in training programs perform poorly in the recognition of normal prepubertal genital anatomy when a labeled photograph is presented (Brayden, Altemieir, Yeager, & Muram, 1991; Ladson, Johnson, & Doty, 1987; Starling, Heisler, Paulson, & Youmans, 2009; Lentsch & Johnson, 2000). Physicians' knowledge of genital anatomy and management of child sexual abuse (CSA) has been found to increase with additional training (Botash et al., 2005; Palusci & McHugh, 1995); however, the ability to correctly interpret the significance of a finding shown in a magnified genital image was still poor (Botash et al., 2005).

In a previous study of examiner agreement (Adams & Wells, 1993), magnified colposcopic images of children's genital and anal findings were shown to 170 medical professionals attending conferences on child abuse between 1990 and 1991. Participants with little experience in performing examinations were significantly more likely to mistake normal variations as signs of abuse. In this study, very little clinical information on the cases was provided to participants. More recently, a study comparing assessments of genital examination findings by Emergency Department physicians with assessments of physicians trained in child abuse evaluation (Makoroff, Brauley, Brandner, Meyers, & Shapiro, 2002) revealed concerning discrepancies. Only 8 of 46 (17%) children identified as having abnormal genital findings, interpreted by Emergency Department physicians as signs of sexual abuse, were found on re-examination by child abuse specialists to actually have abnormal findings. The examination findings of the remaining 38 children were either normal or showed a condition other than abuse, indicating potential errors made by the Emergency Department physicians in both identification and interpretation of genital findings.

Older studies have also reported that when shown photographs of genital findings in girls, clinicians' interpretation of the medical findings changes when a clinical history is provided. In one study, clinicians who regularly examined children for suspected sexual abuse had significantly higher agreement on interpretation of findings in cases when no clinical history was provided ($\kappa = 0.22$) than when history was provided (0.11) (Sinal et al., 1997). In another study, physicians asked to review cases using the same photographs but different clinical histories, on 2 separate occasions, interpreted the photographic findings as being abnormal due to abuse significantly more often when the child's history suggested abuse than when it did not (Paradise, Winter, Finkel, Berenson, & Beiser, 1999). This discrepancy was significantly higher among clinicians with less experience examining children for suspected abuse.

The correct identification and interpretation of genital and anal findings in a child is an important part of the medical evaluation. Correct interpretation involves both recognition of the finding and correlation with the history from the child and other clinical information. Mistaking a normal finding for signs of abuse could lead to an unnecessary investigation, legal consequences, and psychological trauma to the child and family. Training and experience contribute to accurate evaluation of CSA but it is unclear whether there are additional factors which further increase accuracy. Therefore it would be useful to determine the quantity and content of medical training, education, experience, and oversight that is associated with correct identification and interpretation of medical findings in children who may have been sexually abused.

Purpose of study

The present study was designed to:

- (1) Assess medical professionals' ability to recognize normal and abnormal exam findings, ability to interpret the medical and laboratory findings using published guidelines, and to apply knowledge from relevant research.
- (2) Determine which factors in education, clinical experience, and expert review are correlated with greater accuracy in recognition and interpretation of clinical findings.

Methods

Development of the survey instrument

The survey used in this study was pilot tested at two conferences for medical providers in November of 2006 (physicians specializing in child abuse medicine) and January of 2007 (physicians and nurses at a national child abuse conference). Slide presentations were given in a group setting and participants completed a paper questionnaire and answer sheet. Following the presentation, the cases were shown again, giving the answer that was chosen as the correct answer by the presenter (in 2006) and by three experts in 2007. Feedback from participants and review by a panel of expert physicians resulted in the selection of improved photographs for several cases, clarification in the wording of questions, and a clearer format of presentation. The survey was then posted to an online site for this study.

A panel of 10 physician experts was selected based on extensive experience in the medical evaluation of children for suspected sexual abuse (performed over 2,000 examinations of children), routine use of photo-documentation of examination findings, presentations at national conferences, academic productivity, and recognition of expertise by peers through their selection to an honorary society for physician leaders in the field of child abuse medicine (the Ray E. Helfer Society). This panel approved the choice of photographs, wording of questions, and the answers to the survey questions prior to launching the survey.

Three years after the survey was closed to participants, a subset of seven physicians from this original expert group individually reviewed the photographs and answered the survey questions again to confirm the correctness of the responses. All of these experts were subsequently certified by the American Board of Pediatrics as specialists in Child Abuse Pediatrics, following the first board examination in Child Abuse Pediatrics in 2009. Agreement among a minimum of 6 of these seven experts was required to determine the correct answer for each question in consultation with the survey developer. Separate analyses of questions with less than unanimous expert agreement were conducted to assess whether these items were invalid test items or significantly affected respondent total score.

Distribution and content of survey

Using electronic listservs, an invitation to participate in the survey was sent to members of the Ray E. Helfer Society, members of the Section on Child Abuse and Neglect of the American Academy of Pediatrics, members of the Special Interest Group on Child Abuse of the Academic Pediatrics Association, and members of the International Association of Forensic Nurses (IAFN) who perform child sexual abuse examinations. In addition, subscribers to the lists were invited to forward the invitation to colleagues who might also be interested in participating. These groups were selected in order to obtain a convenience sample of medical providers who were likely to be actively involved in the medical evaluation of children for possible sexual abuse.

Volunteer participants were directed, via an electronic link, to a survey posted on the website: <http://www.SurveyMonkey.com> (SurveyMonkey.com, LLC, Palo Alto, CA, USA). Participants completed a questionnaire detailing educational background, specialty, or subspecialty category including self-identification as Child Abuse Pediatricians (CAP), specific education in the field of sexual abuse medical evaluation, experience conducting sexual abuse examinations of children, and clinical practice variables including how often their cases were reviewed by a recognized expert in child sexual abuse medical evaluation. Since board certification through the American Board of Pediatrics was not available at the time of the survey, the identification of specialty as CAP by participants was informal. No identifying information was requested, and the survey was designed so that the Internet (IP) addresses of the responding participants were not stored by SurveyMonkey. A consent form was included as the second page of the survey as required by the Office of Human Research Protection of the University of California, San Diego, who approved the study.

Images and case information for 20 cases of suspected child sexual abuse were presented via the web-based survey, with a total of 40 questions. In 10 of the cases, participants were provided with only the age of the child and the reason for the examination and asked to identify and interpret the findings documented in the de-identified color image of the genital or anal area. No other case information was provided for these questions. In the remaining 10 cases, information in addition to photographs was provided concerning symptoms such as pain, bleeding, itching or discharge, and results of laboratory tests. Seven of these cases also described the child's disclosure of sexual abuse, or denial of abuse, or a description from the child of an accident leading to an injury. Five additional questions were included at the end of the survey to assess participants' recognition of findings from published research studies of children selected for non-abuse, and familiarity with current guidelines from the American Academy of Pediatrics regarding interpretation of sexually transmissible infections in children (Kellogg & Committee on Child Abuse and Neglect, 2005). Each question required a response before the participant was allowed to go to the next question. All questions were multiple-choice or true/false, and in some cases participants were advised when more than one answer could be selected. There were no restrictions given regarding whether or not participants could consult with others before submitting their responses to the survey questions. Comments regarding the questions and answer choices could be entered on the survey as free text.

After the participant answered all the questions on the survey and the test, he/she was instructed to click the "finish" button to submit the responses. After the participants finished the test she/he was directed to a different website for case review with the expert consensus answers highlighted and a discussion of the reasons for the experts' conclusions.

The online survey was open from July 1, 2007 through December 12, 2007. Due to the distribution of the survey through multiple listservs and the suggestion that individuals forward the study link to interested colleagues, it is not possible to determine baseline numbers of professionals who received an invitation to take the online survey. The goal of this study was to solicit responses from specialist Pediatricians, general Pediatricians, nurses, and Advanced Practice Nurses who examine children for suspected sexual abuse and to have enough responses from each group to have sufficient power to detect differences in scores between the groups.

Assessment of validity of questions on survey

Subject responses were entered into a database. Answers chosen by at least 6 of the 7 experts were taken as the correct answer. To assess question validity, a percent correct response rate for all first time takers was calculated and questions

with less than 80% correct were again reviewed to identify those questions that might be confusing or otherwise invalid. In addition, the responses of the 7 experts were analyzed. For 4 questions, fewer than 6 experts chose the same answer, and these questions were eliminated from further analysis. One of these questions asked for a single response but had 2 potential correct answers, two questions concerned disagreements over hymenal rim width measurements, and 1 explored the level of certainty regarding sexual transmission of herpes. Of the remaining 41 questions, 1 point was assigned for each correct response. For questions where more than 1 response could be chosen, the expert panel determined which choices were necessary for the question to be considered correct, which were optional and could vary by practice setting, and which choices were clearly incorrect.

Complete responses were obtained from 197 participants: 118 physicians (60%), 43 Sexual Assault Nurse Examiners (SANE) registered nurses (22%), 33 Advanced Practice Nurses (APN) (17%), 2 nurses (non-SANE), and 1 physician assistant. Approximately 25% of respondents had been involved in child sexual abuse medical evaluation for less than 5 years, 48% for 5–15 years, and 26% for over 15 years.

There were 51 individuals who indicated they had participated or were unsure whether they had participated in a previous pilot survey, presented at a conference for specialist physicians (Ray E. Helfer Society meeting), an interdisciplinary conference for physicians and nurses, or by a self-study course on compact disk. The mean score for these participants was significantly higher than 146 first time participants in the survey (35.4 vs. 31.6, $p < 0.0001$). Repeat takers in all disciplines had higher mean total scores than did first-time takers. Because bias based on previous exposure to the test questions and answers may have raised this group's score, responses from these 51 individuals were excluded from further analysis. Of the remaining 146 first-time survey respondents, 5 did not specify a professional discipline, resulting in 141 first-time takers for whom a professional discipline was reported.

Analyses

Total score results were calculated for 146 first-time participants for the 41 items in the survey which were answered correctly by 80% or more of the respondents using the response chosen by 6 of more of the experts. Total scores were stratified by self-identified respondent discipline and professional specialty and other respondent characteristics. Professional experience (the number of years in this work, number of children ever seen, number of children seen monthly, where children were seen, and number of court appearances) was also stratified by professional discipline.

Questions were then grouped into four categories: (1) identification of the features of anal or genital anatomy depicted in the photographs; (2) interpretation of the significance of findings; (3) clinical decision making; and (4) knowledge of the medical literature. To assess specific areas of competence, additional analysis was done by grouping the questions based on the type of findings shown in the photographs. These categories included normal anatomic variants, accidental trauma, acute and healed genital trauma, and other conditions. Mean scores, standard deviation, median and modal scores were calculated by examiner professional discipline. Student's t tests, comparison of means with Tukey multiple comparisons and regressions were used as appropriate to compare total and subscale scores with training and experience and clinical practice variables in bivariate and multivariate models. Since many of the factors in training, experience, memberships, and journal reading were potentially highly correlated, we first looked at each one individually to see if it was associated with higher or lower score in bivariate analyses and then entered all variables into multiple regression models of total and subscale scores to identify those which were still significant after correcting for potential interactions. Post hoc analyses were used to assess the statistical power of the total and subscale scores to differentiate respondent knowledge levels based on professional discipline and other characteristics.

Results

Background of first time survey participants

Physicians identified themselves as Pediatricians (31) and Child Abuse Pediatricians (39), with the remaining 8 being Family Physicians, Emergency Medicine Physicians, Pediatric Emergency Medicine Physicians, and Gynecologists or other specialists. Nurses identified themselves as Advanced Practice Nurses (APN) (26) or Sexual Assault Nurse Examiners (SANE) (37). The overall mean number of children seen for CSA during their career was 800, indicating that the respondents represented an experienced group of clinicians. The mean number of exams performed was significantly higher among CAP (mean 1310) than SANE (355) and APN (561) ($p = 0.02$). Quarterly or more frequent review of cases and photo-documentation by a recognized expert in child sexual abuse evaluation was provided more for APN and SANE than for Pediatricians and Child Abuse Pediatricians.

We did not collect information about the number of non-abused children or non-abused adults seen by the participant; however, 77% indicated that their current practice included medical care for children with no suspicion of abuse examined in a General Pediatric practice. The mean total score for these respondents was slightly lower than the mean for the group as a whole (0.89 points). Those who responded that they examined patients in other settings had lower mean scores; 5 points lower for participants who examined children in the Pediatric Emergency Department ($p < 0.05$), 7 points lower for practice experience in Gynecology, and 2 points lower for practice in Family Medicine. There were insufficient responses from physicians and nurses working in Gynecology and Family Medicine to show statistical significance of the lower scores.

Table 1

Description of content of survey questions.

<i>Identification of normal anatomy or common variations, from photographs:</i>
Appearance of hymen next to hymen bump/mound
Bump/mound on hymen
Anal fold/part of pectinate line when anus is partially dilated
Venous pooling
Pectinate line in a fully dilated anus
<i>Identification of non-abuse or mimics, from photographs:</i>
Lichen sclerosis
Perianal streptococcal cellulitis
Urethral prolapse
*Perineal groove/failure of midline fusion
*Acute genital injury consistent with accidental mechanism
Labial adhesion
Smooth but relatively narrow rim of hymen
*Lateral notch in hymen, adolescent
Non-significance of anal dilation in prone knee-chest position
<i>Identification and interpretation of abnormal genital findings:</i>
Tear of hymen and posterior fourchette (evidence of penetration)
Acute hymen transection
*Acute hymen tear in a 2 year old could have occurred accidentally
Non-acute hymen transection in an adolescent (evidence of penetration)
<i>Interpretation of sexually transmissible infections:</i>
*Whether to report to protective services a case of perianal condyloma in a 2 year old with likely non-sexual transmission
*Degree of certainty of sexual transmission in 7 year old girl with gonorrhea vulvovaginitis
Use of culture or additional test to confirm Chlamydia infection in child with positive urine NAAT
<i>Knowledge of findings from research studies:</i>
Case control study showed few differences in genital anatomy in girls with and without a history of penetration
Size of hymenal opening not reliable indicator of past penetration
*Lack of case control studies reporting differences in anal dilation in children with and without history of anal penetration
*Knowledge of finding of posterior hymen width measurements of approximately one millimeter in non-abused pre-pubertal girls
*Knowledge of American Academy of Pediatrics recommendations concerning sexually transmissible infections in children for which sexual transmission is certain
*Questions with less than 80% correct response rates

Analysis of question content

Table 1 lists the specific exam findings, medical knowledge, and diagnostic questions that were tested by the questionnaire. Poor test performance was identified when there was lack of recognition that an adolescent hymenal notch was a normal finding, lack of recognition that an acute genital injury may have an accidental etiology, inability to differentiate sexual versus non-sexual transmission of various sexually transmitted infections, and incorrect identification of a perineal groove. Additional questions to assess understanding of research study findings in the published peer-reviewed sexual abuse literature were also used in the assessment of test performance.

The content areas of the questions suggested differences in knowledge among the disciplines. CAP physicians had significantly higher scores compared to SANE nurses regarding correct identification of normal findings in 12 cases (10.5 vs. 9.0, $p < 0.05$) and in 6 cases showing other conditions (5.6 vs. 4.9, $p < 0.05$). For 2 cases of accidental trauma, CAP mean score was significantly higher than APN (1.7 vs. 1.3, $p < 0.05$). There were no significant differences among disciplines for the 6 cases of genital trauma with disclosures from the child of sexual abuse. Questions were also grouped by the amount of clinical information provided, since a previous study (Paradise et al., 1999) indicated that the interpretation of findings on photographs may change with different historical detail. There were 13 questions on cases with no clinical history (Group 1), 5 with minimal history such as symptoms (Group 2), and 18 with information including a disclosure of the child of sexual abuse or accidental injury (Group 3). There was no significant difference in the percent correct scores, however there was a downward trend with increasing amounts of clinical history provided: 81.7% for Group 1, 76.2% for Group 2, and 75.5% for Group 3.

Mean total score for each professional discipline

Fig. 1 shows the distribution of total scores according to discipline, for the lowest range of 0–27, middle range of 28–32, and higher range of 33–41. The distribution of scores across disciplines was significantly different ($p = 0.011$). Significantly more CAP had scores in the higher range compared to Ped (74% vs. 42%, $p = 0.014$) and SANE (75% vs. 35%, $p = 0.0009$), but the difference was not significant comparing CAP to APN (74% vs. 61%, $p = 0.195$). Table 2 shows the mean and standard deviation for total score, the median score, and the modal score for Pediatricians (Ped), APN, SANE, and Child Abuse Pediatricians (CAP). Child Abuse Pediatricians (CAP) had higher mean total scores compared to Pediatricians (Ped) and SANE. Similarly, CAP physicians had significantly higher scores than SANE for 6 knowledge questions (mean 4.6 vs. 3.6), 23 interpretation

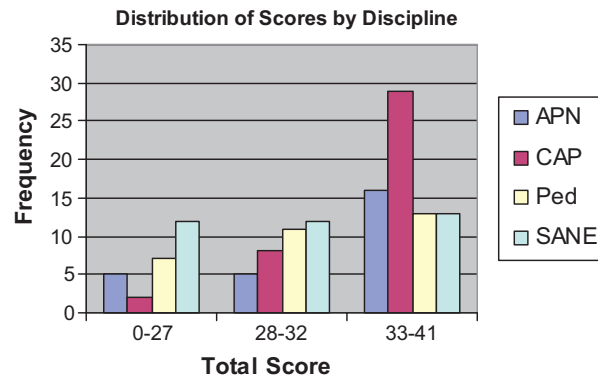


Fig. 1. Distribution of scores by discipline. This graph illustrates the number of participants with total scores in the low range (0–27), mid-range (28–32), and high range (33–41) according to discipline. The proportion of Ped and SANE in the high score range was significantly less than CAP ($p < 0.0001$). APN, Advanced Practice Nurses; CAP, Child Abuse Pediatricians; Ped, Pediatricians; SANE: Sexual Assault Nurse Examiners.

Table 2

Mean total score and distribution by discipline: perfect score = 41.

	N	Mean	SD	Median	Mode
Total group	141	31.6	5.9	33	35
Ped ^a	31	30.1*	7.2	32	32
APN	26	31.5	6.0	33.5	34
SANE	37	29.3*	6.2	28.5	31, 33, and 35
CAP	39	34.8	3.7	35	35

^a Ped, Pediatricians; APN, Advanced Practice Nurses; SANE, Sexual Assault Nurse Examiners; CAP, Child Abuse Pediatricians; SD, standard deviation.

* Difference from CAP, $p < 0.05$.

questions (20.5 vs. 17.9), and 12 clinical decision-making questions (9.7 vs. 7.7), each at $p < 0.05$. However, there were no significant differences in mean subscale scores between CAP physicians and PED, other physicians or APN except with APN in clinical decision-making (9.7 vs. 8.3).

There were sufficient numbers in the CAP, PED, APN, and SANE groups for the 2-way comparison of total score means to have over 80% power to detect a 10% score difference in all comparisons with CAP physicians with $\alpha = 0.05$. Power dropped to below 80% for nurse total score comparisons and for many of the subscale score comparisons in all professional disciplines. Subscale scores were significantly correlated with each other, with correlation coefficients ranging from 0.41 to 0.63, with $p < 0.001$ for all comparisons.

Other factors influencing test performance

All experience, training, self-study, and clinical practice variables were tested in bivariate analysis for effect on total score. Table 3 shows the factors in addition to discipline that were significantly correlated with higher total score. These included certain variables in background, training, experience in child sexual abuse medical evaluations, membership in professional organizations, and type of ongoing self-education in child abuse medicine. Some of these factors did not apply to non-physicians. Membership in the Ray E. Helfer Society is only open to physicians and Child Abuse Pediatrics fellowships are only open to Pediatricians. Both total number of CSA evaluations performed ($p = 0.01$) and average number of evaluations done per month ($p = 0.003$) were significantly associated with a higher total score using bivariate analysis. Since the number of examinations performed monthly was significantly associated with a higher total score, comparison was done to see if there was an effect of professional discipline on scores of participants with higher and lower numbers of monthly examinations performed. Fig. 2 correlates the effect of number of monthly examinations on the mean total score for CAP, Ped, APN, and

Table 3

Experience and activities significantly associated with higher score using bivariate analysis.

Experience	Activity	Contribution ^a
Training	Child Abuse fellowship	+4**
Self study method	Reads <i>The Quarterly Update</i>	+4**
Memberships	Ray E. Helfer Society	+4***
Specialty	Child Abuse Pediatrician	+3.2**
Experience level	Higher total number of child sexual abuse examinations performed	$p = 0.01$
	Higher number of examinations done per month	$p = 0.003$

^a Numbers in parentheses indicate points added to mean score by the factors listed.

** $p < 0.01$.

*** $p < 0.001$.

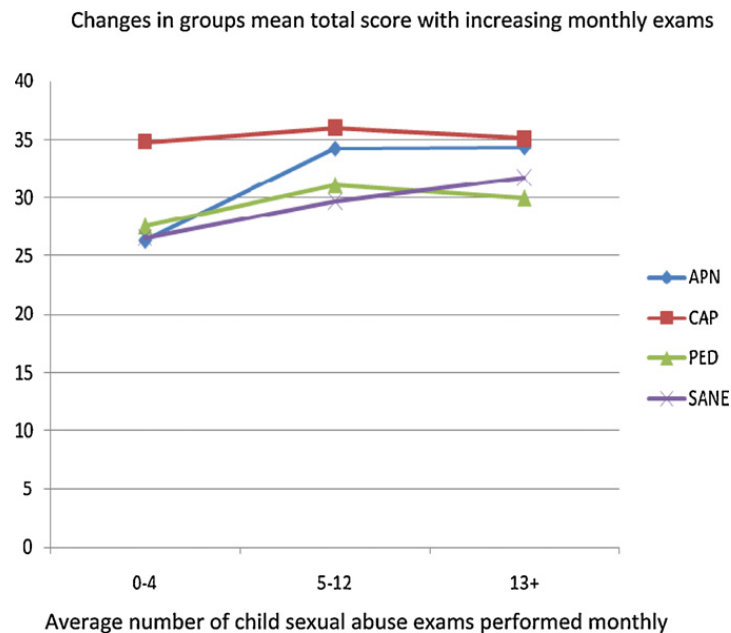


Fig. 2. Total score by discipline and number of cases seen monthly. This graph shows the variation in mean score on the 41 point survey, according to number of sexual abuse examinations performed monthly and professional discipline of the participant. The mean score of non-CAP participants who perform fewer than 5 child sexual abuse exams monthly was significantly lower than CAP participants who examined <5 children monthly ($p < 0.01$). APN, Advanced Practice Nurses; CAP, Child Abuse Pediatricians; Ped, Pediatricians; SANE: Sexual Assault Nurse Examiners.

SANE. Except for CAP, total scores were significantly lower for those participants who examined fewer than 5 children monthly. The mean number of monthly exams was 13.8 in the overall group. There were 42 participants who examined 0–4 children per month for CSA, 41 who examined between 5 and 12 children, and 46 who examined 13 or more children per month, on average. The only significant variation in total score based on number of examinations performed monthly was between non-CAP participants (Ped, APN, SANE) and CAP who examine fewer than 5 children monthly for suspected sexual abuse. Differences in mean score for participants from different disciplines were not statistically significant for those examining “5–12” or “13 or more” children monthly.

Multiple regression analysis of factors associated with higher score

All variables regardless of significance were entered into multiple regression models of the factors associated with higher score controlling for professional discipline. Backward stepwise elimination was used to remove the most insignificant variables one-at-a-time. Table 4 shows the results of this multiple variable regression analysis, while controlling for professional discipline. In this model, only three variables were associated with higher total and subscale scores: (1) reading *The Quarterly Update*, (2) self-identification as a Child Abuse Pediatrician, and (3) review of cases at least quarterly by an expert in child sexual abuse medical evaluation.

The Quarterly Update is a quarterly newsletter summarizing and reviewing published research in the field of child abuse medicine. Reading *The Quarterly Update* was one of the choices in a question asking how participants “keep up to date with the advances in research in child sexual abuse medical evaluation”. Other choices were “attend conferences”, “read medical journals on a regular basis”, “read journal abstracts”, or through “newsletters”. Reading *The Quarterly Update* raised scores

Table 4
Factors associated with higher total score (multivariable regression models).

Factor	Contribution to total mean score (P) ^a	Contribution to knowledge score (P) ^a	Contribution to interpretation score (P) ^a	Contribution to clinical decision score (P) ^a
Expert review of cases done at least quarterly	+3.0**	+0.68***	+1.5**	ns
Child Abuse Pediatrician	+2.9*	ns	ns	1.1**
Reads <i>The Quarterly Update</i>	+3.8***	+0.71***	+2.6***	0.88*
Regression F , adjusted R^2	<0.0001, 0.2331	<0.0001, 0.1383	<0.0001, 0.1638	<0.0001, 0.1336

^a Numbers indicate points added to mean score, associated with examinee attribute.

* $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$.

ns: $p > 0.05$.

5–10%, while controlling for other significant factors in the model. None of the other methods for keeping up to date were significantly associated with higher scores.

Two factors that were significantly associated with higher score in the bivariate analysis remained significant after multiple regression analysis. Self-identification as a Child Abuse Pediatrician raised both total score and clinical decision subscale score. Reading *The Quarterly Update* significantly raised total score and all three subscale scores. Having cases reviewed at least quarterly by an expert in CSA evaluation, using photo-documentation, raised total score and knowledge and interpretation subscale scores. This practice variable was significantly correlated with better test performance using the multivariable regression model, controlling for professional discipline.

In these models, there was no significant association with the number of years since first training was obtained, the total number of sexual abuse examinations performed, the number of monthly examinations performed, the frequency of attendance at conferences on sexual abuse, experience lecturing at regional, national, or international conferences, academic productivity, defense testimony, or the number of times testifying in court.

Discussion

Comparison with other research findings

Similar to a previous study using colposcopic photographs (Adams & Wells, 1993), higher scores were significantly associated with both the total number of sexual abuse evaluations performed ($p = 0.01$) and the average number of evaluations performed monthly ($p = 0.004$). Our study identified additional factors that were independently correlated with higher scores: self-identification as a Child Abuse Pediatrician, having cases reviewed by an expert in sexual abuse evaluation at least quarterly, and reading *The Quarterly Update* publication as a means of keeping up to date with advances in research in the field. In 2006 and 2007, the percentage of articles in *The Quarterly Update* related to CSA ranged from 5% to 10% per issue, as counted by the first author. Reading reviews in this publication appears to be a proxy for keeping up with new research in the field of child abuse in general. While anyone may subscribe to *The Quarterly Update*, it is likely that most subscribers are physicians.

Survey participants who self-identified as Pediatric Emergency Medicine Physicians had a mean total score that was significantly lower than the overall group of first-time takers. This finding is similar to the Makoroff study (Makoroff et al., 2002), which showed very low agreement (17%) between Pediatric Emergency Medicine Physicians and child abuse pediatric specialists on the identification and interpretation of physical findings suggestive of sexual abuse.

Our study included a higher percentage of non-physician participants: 47% of the total group compared to 7% in the study conducted in 1990 and 1991 (Adams & Wells, 1993), which may reflect the change in practice in many communities where more APN or SANE now perform child sexual abuse medical evaluations. Because there were sufficient numbers of survey participants from different professional disciplines, appropriately powered comparisons could be made between groups in this study for total scores.

Review of cases at least quarterly “by a recognized expert in child sexual abuse medical evaluation”, using photo-documentation, significantly raised total score after controlling for professional discipline. This is a practice that has not been previously identified as improving diagnostic accuracy in CSA examination. Non-specialist Pediatricians, Advanced Practice Nurses, and Sexual Assault Nurse Examiners as well as Child Abuse Pediatricians can now participate in ongoing expert case review through such organizations as the Telehealth Institute for Child Maltreatment (THICM). THICM provides case reviews by Child Abuse Pediatrics specialists in an anonymous online format as a method of quality improvement for providers of CSA medical evaluations (www.thicm.org). Other methods of expert case review, both anonymous and direct, are available depending on local resources.

Limitations

Response rate

Our study had several limitations. Those individuals who responded to the invitation to participate in the survey represent a convenience sample of health care professionals who examine children for suspected sexual abuse, and cannot be said to be a representative sample of any of the professional disciplines involved. There is no national directory of physicians or nurses who examine children for possible sexual abuse, so it was not possible to send the survey to a select group and determine the response rate. We were not able to determine whether more or less experienced clinicians were more likely to complete the survey, but the mean number of CSA examinations performed by first time participants was 800, indicating that it was a fairly experienced group.

Survey design

This study used still images of genital and anal findings rather than video recordings that capture the dynamic nature of the examination. Being able to view an entire examination on video may provide a more complete picture of the tissues and how they change with different types of examination techniques (Palusci, Cox, Shatz, Wakefield, & Buchanan, 1998). It

is difficult to be certain looking only at still photographs that the finding seen in the photograph is a true representation of the appearance of the genital or anal anatomy; this was mentioned in comments from some of the participants who wanted additional views presented before reaching a conclusion.

The ability of participants to answer the questions, submit their survey, and then have access to a site where the answers were presented could have resulted in falsely elevated scores, although there was no indication that this had actually occurred. Some of the questions included making a diagnosis, a skill that is outside the scope of practice of Sexual Assault Nurse Examiners. However, subscale scores for SANE on identification of normal findings and conditions other than abuse were also lower than Child Abuse Pediatricians.

Scores on an examination such as this may not reflect actual clinical proficiency. An online survey is a poor approximation of clinical practice. Participants who regularly have their cases reviewed may have, in actual practice, the opportunity to obtain a second opinion from another member of the medical team before reaching a conclusion about the identification and interpretation of an unusual medical finding. The use of an individual online survey format does not incorporate this possibility.

In addition, many of the factors used to measure examiner experience were inter-related in that more experienced participants also used more ways to keep up to date and had more court and ongoing clinical experiences. These likely reduced the power of our analyses to detect other important contributions to total and subscale scores.

Emphasis on physical findings rather than history from the child

This study emphasized identification and interpretation of findings in brief case summaries, in some cases eliminating much of the history and ancillary case data. The absence of the complete historical information in several cases was problematic for some participants and this was reflected in the comments section of the survey. Other comments from participants referred to the need for additional views or improved quality of the photographs. Several participants commented that protocols in their particular practice setting dictated whether reports to protective services should be made for cases of possible non-sexual transmission of some sexually transmissible infections.

The diagnosis of sexual abuse, of course, involves much more than interpreting medical and laboratory findings, but this study did not address those other important factors. It is well recognized that the child's description of the abuse (Kellogg & Committee on Child Abuse and Neglect, 2005) and any report to the medical provider of symptoms experienced at the time of the incident (DeLago, Deblinger, Schroeder, & Finkel, 2008) provide additional critical evidence and information in the overall evaluation of the suspected abuse.

The diagnosis of any disease or condition requires the integration of medical history, physical examination findings and laboratory results. For some diseases the examination findings are so distinct they are pathognomonic or diagnostic. Many diseases have similar presenting signs and symptoms, and clarity is obtained through laboratory testing. In cases of child sexual abuse the medical history is the critical piece to understanding what a child may have experienced. The medical history is paramount because very few children present with diagnostic examination findings. Children are capable of providing detailed, age-appropriate and idiosyncratic statements that enrich our understanding of what they experienced and help explain the presence or absence of examination findings. This study reinforces the importance of the correct interpretation of physical findings, but we stress that these findings do not exist in isolation. The clinician has an obligation to obtain detailed medical histories in a developmentally appropriate manner that is non-leading, not suggestive, and non-judgmental. A correctly obtained history integrated with the physical examination findings and laboratory tests provides an opportunity to formulate an accurate (and legally defensible) diagnosis.

On the other hand, a history by the child of being sexually abused may also influence less experienced physicians or nurses to over-interpret other conditions or examination findings that are variants of normal as being signs of abuse. Other researchers have found this to be the case in previous studies using both photographs of medical findings and historical information (Paradise et al., 1999; Sinal et al., 1997).

Significance and medico-legal implications

The use of total score and subscale score in our study was an attempt to quantify participants' knowledge of medical research findings and ability to recognize and interpret the significance of findings from photographs. In a clinical setting, mistaking a normal anatomic variant, a condition other than abuse, or an injury caused by an accident as being due to sexual abuse has serious adverse consequences. Even one mistake of this type can be devastating for the child and family. Physicians not specializing in child abuse evaluation, Advanced Practice Nurses and Sexual Assault Nurse Examiners who examine fewer than five children monthly for suspected sexual abuse all performed poorly on this survey. For these less experienced providers, review of every case by an expert in child sexual abuse evaluation could decrease the likelihood of a misdiagnosis of child abuse.

Conclusions

This study sought to determine the ability of survey participants to correctly identify genital and anal findings and interpret medical findings using knowledge of the literature, and to understand the significance of sexually transmissible

infections. Both the mean total score and subscale scores on the survey were significantly higher among participants who self-identified as Child Abuse Pediatricians (CAP) compared to those who self-identified as Sexual Assault Nurse Examiners (SANE) and others. There were no statistically significant differences between CAP and Advanced Practice Nurses (APN) or between Pediatricians and SANE in some measures. Training, discipline, and clinical experience were significantly associated with the ability to correctly identify medical findings and apply medical knowledge to correctly interpret findings in our survey of participants who perform evaluations for suspected child sexual abuse. Beyond these, ongoing practice, expert case review, and keeping up with the medical literature appear to confer additional diagnostic accuracy. Among the 141 first time participants, there was a direct correlation between the number of exams performed monthly and a higher mean score, but a significant difference was noted between non-CAP and CAP participants examining fewer than 5 patients monthly for suspected CSA. Our data suggest that 5 or more examinations per month may be required for ongoing competency in interpreting medical and laboratory findings in children evaluated for suspected sexual abuse by clinicians other than those specializing in Child Abuse Pediatrics. Results from this study suggest that the cases of examiners not meeting these criteria should be reviewed by a more experienced specialist medical provider to improve the accuracy of CSA medical evaluations.

The recognition of injuries as well as normal variants and conditions confused with abuse is critical to formulating an accurate diagnosis or conclusion when children are evaluated for suspected child sexual abuse. Additional studies are needed to develop methods of assessing examiner competence to ensure that children evaluated for suspected sexual abuse receive competent and comprehensive diagnostic and therapeutic care.

References

- Adams, J. A., & Wells, R. (1993). Normal versus abnormal genital findings in children: How well do examiners agree? *Child Abuse & Neglect*, 17, 663–675.
- Botash, A. S., Galloway, A. E., Booth, T., Ploutz-Snyder, R., Hoffman-Rosenfeld, J., & Cahill, L. (2005). Continuing medical education in child sexual abuse: Cognitive gains but not expertise. *Archives of Pediatric and Adolescent Medicine*, 159(6), 561–566.
- Brayden, R., Altemeier, W., 3rd, Yeager, T., & Muram, D. (1991). Interpretations of colposcopic photographs: Evidence for competence in assessing sexual abuse? *Child Abuse & Neglect*, 15(1–2), 69–76.
- DeLago, C., Deblinger, E. D., Schroeder, C., & Finkel, M. A. (2008). Girls who disclose sexual abuse: Urogenital symptoms and signs following sexual contact. *Pediatrics*, 122(8), e221–e226.
- Kellogg, N., & Committee on Child Abuse and Neglect, American Academy of Pediatrics. (2005). The evaluation of sexual abuse in children, clinical report. *Pediatrics*, 116(2), 506–512.
- Ladson, S., Johnson, C. F., & Doty, R. E. (1987). Do physicians recognize sexual abuse? *American Journal of Diseases in Children*, 141(4), 411–415.
- Lentsch, K. A., & Johnson, C. F. (2000). Do physicians have adequate knowledge of child sexual abuse? The results of two surveys of practicing physicians, 1986 and 1996. *Child Maltreatment*, 5(1), 72–78.
- Makoroff, K. L., Brauley, J. L., Brandner, A. M., Meyers, P. A., & Shapiro, R. A. (2002). Genital examinations for alleged sexual abuse of prepubertal girls: Findings by pediatric emergency medicine physicians compared with child abuse trained physicians. *Child Abuse & Neglect*, 26(12), 1235–1242.
- Palusci, V. J., Cox, E. O., Shatz, E. G., Wakefield, S., & Buchanan, G. (1998). *Let's go to the videotape: Comparing video and still images in the assessment of child sexual abuse*. Presented at the Annual meeting of the Ambulatory Pediatric Association, New Orleans, LA, May 4.
- Palusci, V. J., & McHugh, M. T. (1995). Interdisciplinary training in the evaluation of child sexual abuse. *Child Abuse & Neglect*, 19(9), 1031–1038.
- Paradise, J. E., Winter, M. R., Finkel, M. A., Berenson, A. B., & Beiser, A. S. (1999). Influence of history on physicians' interpretation of girls genital findings. *Pediatrics*, 103(5), 980–986.
- Sinal, S. H., Lawless, M. R., Rainey, D. Y., Everett, V. C., Runyan, D. K., Frothingham, T., Herman-Giddens, M., & St. Claire, K. (1997). Clinician agreement on physical findings in child sexual abuse cases. *Archives of Pediatric and Adolescent Medicine*, 151(5), 497–501.
- Starling, S. P., Heisler, K. W., Paulson, J. F., & Youmans, E. (2009). Child abuse training and knowledge: A national survey of Emergency Medicine, Family Medicine, and Pediatric residents and program directors. *Pediatrics*, 123(4), e595–e602.