

▲ An Investigation of Interrater Reliability Among Athletic Training Accreditation Site Visitors

Jeffrey W. Wimer, PhD, ATC

Accreditation plays an important role in many allied health professions. The main purpose of this study was to determine interrater reliability among athletic training accreditation site visitors assigned to interpret and judge a college or university's educational program in accordance with published standards. Ten scenarios were sent electronically to 135 site visitors, of which 93 (68%) responded. Respondents rated whether hypothetical situations described in each scenario were compliant (i.e., meets standard) or non-compliant (i.e., does not meet standard). A comment section was included for site visitors to justify their responses and/or to clarify their selections. The results of this study showed that there was poor interrater reliability in the judgments made by the site visitors. Moreover, a majority of the respondents responded to two of the 10 scenarios incorrectly (scenarios 5 and 10). The average score was 6.77 ± 1.57 , or 68%. Site visitors who misinterpret standards or apply personal values pose serious threats to the accreditation process because they may judge educational programs compliant when in fact programs are noncompliant and vice versa. Furthermore, this study suggests that there may be a great deal of variability across academic programs in applying accreditation standards because of the low agreement among site visitors. Therefore, educational methods and training procedures for accreditation site visitors should be continually reexamined in an attempt to improve and ensure consistency in allied health professions. *J Allied Health* 2005; 34:65-75.

ACCREDITATION STANDARDS are driving curriculum development and assessment strategies on college campuses across the country. However, accreditation is not new. More than a century ago, accreditation agencies formed as self-regulatory and evaluative forces for higher education in the

Dr. Wimer is Assistant Professor and Athletic Training Program Coordinator, Department of Wellness and Sport Sciences, Millersville University, Millersville, Pennsylvania.

Received May 12, 2003; revision accepted April 28, 2004.

Address correspondence and reprint requests to: Jeffrey W. Wimer, PhD, ATC, Department of Wellness and Sport Sciences, Millersville University, P.O. Box 1002, 1 South George Street, Millersville, PA 17551-0302. Tel 717-871-5367; fax 717-871-2393; e-mail jeffrey.wimer@millersville.edu.

United States.¹⁻⁶ Today, there are regional and specialized accreditation agencies. Regional agencies grant institutional accreditation to an entire college or university, whereas specialized accreditation agencies such as the National League for Nursing Accrediting Commission, the Commission on Dental Accreditation, and the Commission on Accreditation of Allied Health Education Programs (CAAHEP) grant accreditation to specific educational programs.

Across the United States, there are six regional accreditation agencies representing different geographic locations (i.e., New England, Middle States, North Central, Southern, Northwest, and Western). The primary goal of regional accreditation is to ensure quality teaching and learning vis-à-vis established and accepted standards.⁷ Specialized agencies vary widely and represent various professions. For instance, CAAHEP reviews and accredits more than 2,000 educational programs in 21 health science occupations such as emergency medical technician/paramedic, orthotist/prosthetist, respiratory therapist, and surgical technologist.⁸

The concept of specialized accreditation for athletic training education programs has been important in the preparation of certified athletic trainers (ATC) for more than 40 years.⁹⁻¹¹ Accreditation standards provide a comprehensive framework in which the clinical instruction of students is monitored. Program directors and clinical instructors rely on accreditation standards to help them maintain stability when faced with budget cuts or declining student enrollments. In addition, accreditation helps to bring legitimacy to the profession, especially as ATCs seek third-party reimbursement for their services from medical insurance companies and health maintenance organizations.¹² Only accredited programs may sponsor students for the National Athletic Trainers' Association Board of Certification (NATABOC) examination. Accreditation is indeed an important strategy for assuring quality control not only in athletic training but also in many allied health professions.

For the past 10 years, academic programs in athletic training have been accredited by CAAHEP. Recently, the athletic training profession dissociated from CAAHEP in part because the agency was perceived as too restrictive for future growth.^{13,14} The agency currently responsible for the education and training of athletic training site visitors is the Joint Review Committee on Educational Programs in Athletic Training (JRC-AT). This organization is cosponsored by the National Athletic Trainers' Association

(NATA), American Academy of Family Physicians, American Academy of Pediatrics, and American Orthopaedic Society for Sports Medicine. The JRC-AT is transitioning itself to become the sole accreditation agency for approximately 300 institutions that sponsor academic programs to prepare athletic trainers.

Unfortunately, the process of obtaining accreditation from specialized agencies like CAAHEP and the JRC-AT has come under attack similar to the ways regional accreditation agencies experience criticism.^{13,15} Some academics believe accreditation agencies are closed circles of people who make decisions in secret.¹⁶ Others in the higher education arena argue the process is too costly, suggesting the agencies are ineffective bureaucracies that threaten institutional autonomy and thus ought to be voluntary rather than mandatory.¹⁷⁻²⁰ Part of the skepticism, it seems, occurs when decisions about compliance with standards vary among site visitors. According to former CAAHEP president and certified athletic trainer Larry Leverenz, PhD, ATC, "Many of the problems we hear regarding accreditation and the process have to do with the site visitors. The site visitor is also the link in the accreditation chain that we have the least control over. It's an interesting phenomenon" (personal communication, August 8, 2001). Because humans, not machines, are applying judgment, allied health professionals and their accreditation organizations must develop mechanisms to ensure fairness. In other words, the people assigned the task of evaluating academic programs must be able to assess accurately and derive similar meanings from the same observable event.

One method for ensuring fairness among accreditation evaluators is to develop precise guidelines to use in assessing compliance with standards. In athletic training, for example, an interpretation manual for the 2001 CAAHEP accreditation standards and guidelines²¹ was developed to assist in the collective understanding of compliance and noncompliance issues within this profession. At present, however, there are no mechanisms in place to ensure that all site visitors have read and accurately interpreted the manual. Therefore, this study provides a method for understanding how site visitors in athletic training make decisions regarding compliance with accreditation standards.

At present, site visitor training occurs when interested candidates send their curriculum vita and cover letter to the JRC-AT. The JRC-AT chairman screens top candidates and invites a cadre of potential site visitors to a training session that typically coincides with an event such as the annual NATA Educators Conference or the annual NATA meeting and clinical symposium. During the workshop, various topics are discussed with candidates, including the structure and function of the JRC-AT, review of the standards and guidelines, self-study documents, the on-site visit, the on-site visit report, scheduling and administrative procedures, and conflict-of-interest scenarios.²² A written evaluation is distributed at the conclusion of the workshop. New site vis-

itors are paired with more experienced site visitors in order to mentor the candidates during their initial visits.

The accreditation site visit and the individuals selected to conduct on-site evaluations have not been examined extensively in the allied health literature. Therefore, the major research questions asked in this study were as follows. Is there interrater reliability among site visitors? Are site visitors accurate in their assessments of scenarios? In what ways do site visitors justify their compliance and noncompliance decisions?

Methods

DESIGN

As suggested by the title of this article, the main purpose of this study was to determine the level of interrater reliability among athletic training accreditation site visitors in the interpretation and judgment of the 2001 CAAHEP accreditation standards and guidelines. To determine this, an instrument²³ consisting of 10 scenarios describing situations encountered during on-site visits was developed to investigate site visitor responses under identical conditions. Two quantitative measures were used to calculate responses: the intraclass correlation coefficient (ICC) statistic and the χ^2 statistic.

A secondary purpose of this study was to investigate whether site visitors were accurate in their assessments of the scenarios. This was determined by calculating an average score for identifying the scenarios correctly. In other words, one point was given for each correct response and zero for incorrect responses. The total number of possible points respondents could receive was 10. The total number of points earned was converted into the percent correct and incorrect.

Using key words and phrases taken directly from the interpretation manual for the 2001 CAAHEP accreditation standards and guidelines,²¹ the 10 scenarios in the study were developed to represent various situations a site visitor might encounter during an actual on-site visit. Much like a self-study document, the scenarios contained various facts such as student enrollment, regional location, and the leadership characteristics of key officials associated with the school.

Drawing upon the researcher's experiences as a site visitor, accreditation consultant, and the program director of a CAAHEP-accredited athletic training education program, seven of the scenarios represented actual situations previously encountered during the past 10 years. The remaining three scenarios were developed from conversations and discussions with other site visitors. Fictitious college and university names were used to eliminate any identifying information. A list of the standards and scenarios used in this study is located in the Appendix.

To investigate any potential demographic differences and the possibility that a particular variable might influence or account for judgment ratings, respondents were divided into several demographic categories such as gender, highest degree earned, NATABOC route to certification,

political view, and seasoned or novice. A point biserial correlation coefficient statistic, which assesses the degree of association between a dichotomous variable and an interval/ratio variable, was calculated for the amount of preparation site visitors spent reading self-studies and studying documents before conducting an actual site visit.

An additional purpose of this study was to understand the underlying meaning of the decision-making process of site visitors in relation to each scenario. More specifically, in what ways do site visitors justify their compliance and non-compliance decisions related to each scenario? To explore this question, a qualitative method suggested by Lincoln and Guba²⁴ was used to give insights about the qualitative dimensions of agreement and disagreement related to site visitor justifications and to give meaning to the lack of agreement. This evaluation was determined through analysis of the open-ended portion of the instrument whereby individuals could comment on their response.

INSTRUMENT

Because the scenarios produced were at the heart of this study, several steps were taken to ensure they were valid and reliable representations of the standards that each was meant to exemplify. First, a statistical consultant with a doctorate in experimental design and statistics was hired to evaluate the research procedures and to approve the statistical analysis techniques. To assess construct validity and reliability, five experts in athletic training accreditation were selected to judge the degree to which each scenario represented the accreditation standard. The expert panel included (1) the former president of CAAHEP, (2) the current chairperson of the JRC-AT, and (3) three experienced site visitors, each of whom had served as the director of their own nationally accredited educational program in athletic training.

Each expert was asked to provide suggestions for appropriate choice of wording of scenarios (familiarity, difficulty, and grammar) and to review the online instrument for content and face validity. Minor editorial changes were recommended for the directions and readability of the scenarios, and the instrument was revised according to these recommendations.

The reliability of the total instrument was estimated by using the Kuder-Richardson formula (KR21). The KR21 functions on the premise that items on the instrument are equally difficult. Because this assumption is seldom met, the formula is considered a conservative estimate of reliability.²⁵ In accordance with the recommendations of Morrow et al.,²⁶ the goal of this investigation was to achieve a reliability of 0.80 or higher. The actual reliability for the instrument was calculated to be 0.86. Panel responses were also compared with responses provided by Danny Foster, PhD, ATC, the lead author of the interpretation manual for the 2001 CAAHEP accreditation standards and guidelines.²¹ The expert panel results matched the results of those provided by the lead author of the

interpretation manual. As an additional measure, the JRC-AT executive committee, which consists of eight ATCs, approved the research methodology for this study before implementation.

PROCEDURES

In January 2002, all athletic training accreditation site visitors were sent two e-mail memorandums written by the chairperson of the JRC-AT informing them about the study and the link to the instrument Web site. After accessing the instrument, respondents read a list of 10 published standards from the 2001 CAAHEP accreditation standards and guidelines followed by 10 corresponding scenarios encountered in the field. Respondents rated whether the education program described in each scenario was compliant with the standard listed (i.e., meets standard) or noncompliant (i.e., does not meet standard) by clicking on a pull-down box marked "compliance" or "noncompliance" after each scenario. A comment section in the form of a text box was included under each pull-down box for respondents to justify their responses. Following the scenarios, respondents were provided with a brief survey of demographic information such as gender, highest degree earned, political view, and years of experience in the profession. Participation in this study was voluntary, and anonymity was assured. Consent was described in the directions and stated to be implied upon the respondent's submission of his or her responses by clicking the "submit" button at the end of the instrument. Clicking the "submit" button automatically transferred the raw data to a computer file at the researcher's university. Upon clicking the "submit button," respondents were sent an automatic pop-up window acknowledging their participation.

SUBJECTS

Ninety-three of the 135 (68.8%) site visitors (61 men and 32 women; age, 41.1 ± 7.6 years; ATC experience, 17.7 ± 6.6 years) participated in the study (personal communication, L. Caruthers, December 18, 2001). These respondents reported they typically spent more than 20 hours (21.5 ± 12.3 hours) reviewing documents and preparing for a site visit. Between them, the site visitors conducted 6.6 ± 8.3 visits and served as team chair 3.1 ± 6.2 times. Although all members of the population had completed a training workshop to become an evaluator, some of the respondents either had yet to perform a visit ($n = 12$; 16.7%) or did not respond to the question regarding number of visits conducted ($n = 21$; 22.5%). Table 1 includes subject demographic characteristics.

DATA ANALYSIS

To assess interrater reliability, two quantitative measures were used: the ICC statistic²⁷ and the χ^2 statistic.²⁸ ICC values greater than ± 0.75 represent excellent agreement

TABLE 1. Demographic Characteristics of Respondents

Characteristic	n	%
Response rate	93	(68.8)
Gender		
Male	61	(65.6)
Female	32	(34.4)
Highest degree earned		
Doctoral degree	65	(69.9)
Master's degree	28	(30.1)
Route to NATABOC certification		
Curriculum	61	(67.0)
Internship	30	(33.0)
Political view		
Far left	0	(0)
Liberal	12	(13.8)
Middle of the road	43	(49.4)
Conservative	31	(35.6)
Far right	1	(1.1)

Note. $n = 135$ (68.8%).

NATABOC, National Athletic Trainers' Association Board of Certification.

beyond chance, values of ± 0.40 – 0.75 represent fair to good agreement beyond chance, and values less than ± 0.40 represent poor or no agreement.²⁷ The χ^2 statistic, a nonparametric statistic, was used to test whether the majority of respondents correctly identified the scenarios as compliant or noncompliant. All statistical analyses were considered significant at the $p < 0.05$ level, and statistics were computed using SPSS.²⁹

To determine whether reviewers were accurate in their assessments of the scenarios, an average score was calculated for each reviewer. In other words, accuracy was determined by the percent correct and incorrect, that is, one point was given for each correct response and zero for incorrect responses. Thus, the total number of possible points respondents could receive was 10.

To identify the ways in which site visitors justify their compliance and noncompliance decisions related to each scenario and to better understand the underlying meaning of the decision-making process, a qualitative method suggested by Lincoln and Guba²⁴ was used. The qualitative method entailed data reduction and data display in order to develop the themes from each of the scenario responses. Likewise, several steps were taken to increase trustworthiness in this process. For instance, an audit trail was carefully maintained by printing hard copies of each respondent's online submission for external review. Triangulation was accomplished by reading the respondents' comments aloud to two colleagues to authenticate their meaning. Later, the frequency of each emerging theme was recorded after each unit of data had been examined for meaning and categorized into like categories.

For purposes of data management, comments chronicled less than three times on the researcher code sheet were not

TABLE 2. Reliability Analyses

Subjects	ICC	ICC 95% Confidence Intervals		<i>p</i>
		Lower Limit	Upper Limit	
Site visitors	0.369	0.153	0.549	0.001

ICC, intraclass correlation coefficients ($p < 0.05$).

reported in the results, whereas comments appearing more than three times were established a priori to be themes. This is not to say single comments were not important; rather, they were deemed minority comments for this study. Furthermore, several respondents had no comment. Therefore, the actual number of prevailing themes does not completely represent the totality of comments from the respondents. For example, a single comment submitted by one respondent was not recorded unless the same comment was recorded at least three discrete times during the coding process to become a theme.

Results

Based on this sample, interrater reliability among the athletic training site visitors was low. The majority of respondents answered scenario 5 and scenario 10 incorrectly. Across the 10 scenarios, the ICC statistic suggests site visitors do not rate compliance and noncompliance similarly (Table 2). The ICC was 0.369 with a confidence interval of 0.153–0.549. Although the coefficient is statistically significant and the respondents performed better than chance on a majority of the scenarios (Table 3), their value of agreement is considered poor to fair.²⁷ The average accuracy score for identifying the scenarios correctly as compliant and noncompliant was 6.77 ± 1.57 (68%).

χ^2 analyses also showed no significant difference in ratings of compliance and noncompliance between seasoned site visitors versus novices. The χ^2 values ranged from 0.006 to 2.55. As per JRC-AT practices, seasoned was defined as three or more visits as team chair. Along gender lines, there was no significant difference in ratings except scenario 6, in which female site visitors were more in agreement than male site visitors ($p = 0.020$), rating the scenario noncompliant. Regarding preparation time, no significant differences were found between site visitor ratings except in scenario 6. As preparation time increased, site visitors were more likely to judge scenario 6 noncompliant ($r = 0.28$, $p = 0.039$). There was no significant difference in the number of hours men and women spent preparing for site visits ($t = -0.282$, $p = 0.779$).

Political views held by site visitors appeared to have no bearing on ratings of compliance and noncompliance in this study. Personal political ideology, however, has been shown to influence professional practice in other fields, namely clinical and counseling psychology.³⁰ More impor-

tantly, it has been demonstrated that a person's location on a liberal-conservative continuum can affect the decisions he or she may render and the recommendations he or she provide.³¹ Therefore, it is reasonable to conclude that one's personal beliefs could influence decision making. To test the influence of political view on judgments of compliance, the first step was to identify the percentage of site visitors who identified with liberal and conservative values. The next step was to investigate if self-reported political ideology related to scenario judgments. Several χ^2 analyses revealed, however, no significant differences between the site visitors' political view and their ratings of compliance. The χ^2 values ranged from 0.791 to 3.91, and none were statistically significant.

There were no significant differences in ratings of compliance and noncompliance between site visitors with master's versus doctoral degrees (χ^2 values ranged from 0.014 to 1.69). An individual's route to NATABOC certification had no impact on ratings as well (χ^2 values ranged from 0.090 to 3.07). In short, demographic factors such as gender, experience, political view, time spent preparing, degree, and route to certification did not appear to influence respondents' judgment ratings in this study.

Qualitative analysis of the comments provided by the respondents suggests that several site visitors used faulty logic or applied personal values when rating items compliant or noncompliant (Table 4). For example, in scenario 1, several site visitors clarified their remarks by stating they believed the scenario referred to learning over time and the documentation of skill mastery. Although there is some question as to student competency in clinical proficiency assessment in scenario 1, the main purpose was to determine each respondent's thinking regarding isokinetic equipment owned by an off-campus facility. Although clinical proficiency is important, that is not what is being measured by this standard; rather, what is being measured is the students' regular exposure to the equipment. Interestingly, at least four site visitors questioned the inappropriateness of isokinetic equipment as an accreditation standard. As one site visitor wrote, "Someday we [as a profession] have to ask why we teach some of the things we do. The isokinetic dynamometer [sic] is becoming a dinosaur in many practices." This respondent rated the scenario incorrectly, which may imply an individual bias.

In scenario 2, several site visitors indicated that the institution described in the essay should have developed methods for evaluating each student's psychological status. The inclusion of students with various disabilities has indeed changed the climate in higher education, and a student's mental status is indeed an important priority; however, institutions are not legally responsible for conducting psychological screenings before admitting students. Students must be informed and verify that they meet the technical standards required for the athletic training profession, but they do not need to submit to psychological testing before admission. This information, including the proper

TABLE 3. Percent Compliant and Noncompliant

Scenario	Compliant (%)	Noncompliant (%)	χ^2	<i>p</i>
1	74.2*	25.8	21.8	0.000
2	31.2	68.8*	13.2	0.000
3	25.8	74.2*	21.8	0.000
4	80.6*	19.4	34.9	0.000
5	43.5*	56.5	1.6	0.211
6	20.0	80.0*	32.4	0.000
7	70.7*	29.3	15.7	0.000
8	80.6*	19.4	34.9	0.000
9	30.1	69.9*	14.7	0.000
10	60.4	39.6*	4.0	0.046

*Correct response, *p* < 0.05.

procedures for handling such situations, was available in the interpretation manual for the 2001 CAAHEP accreditation standards and guidelines.²¹

Scenario 3 suggests that there is only one method for evaluating students when in fact educational programs must demonstrate multiple assessment methods. The educational program described in this scenario does not allow students to benefit from early and perhaps frequent feedback for improvement, and therefore this scenario was noncompliant. Many site visitors recommended that students be evaluated at the middle and end of the term, which is correct.

In scenario 4, institutions have the autonomy to develop the sequence of courses if appropriate steps are taken to determine the adequacy of the progression and evidence supports an enriching overall educational experience. Without reading the course descriptions and syllabi, rating this scenario noncompliant may be inappropriate because some kinesiology courses focus on advanced topics such as biomechanics and motion analysis rather than muscle origins, insertions, and actions and therefore would be unsuitable for students to take before therapeutic exercise. The themes expressed by site visitors in this scenario, however, varied from calls for improved sequencing, which is incorrect, to calls for improved institutional autonomy, which is acceptable.

In scenario 5, the program is using data to make programmatic changes. The frequency of administration of such surveys is not defined in the standards or in the interpretation manual and therefore is considered compliant. Although several respondents noted the need for more frequent use of both quantitative and qualitative methodologies, which is an appropriate recommendation, the majority of respondents judged this scenario incorrectly, instead insisting that programs conduct yearly data-driven evaluations.

In scenario 6, the timing of calibrations for electronic modalities and machinery is clearly defined in the interpretation manual for the 2001 CAAHEP accreditation standards and guidelines.²¹ Therefore, this scenario is considered noncompliant because the program elected to postpone formal calibration and safety checks of modalities. Many respondents believed that patient safety was in jeop-

TABLE 4. Frequency of Themes Provided by Site Visitors

Scenario	No.	Prevalent Themes	Scenario	No.	Prevalent Themes
1	10	Documentation meets intent of standard	6	26	Patient safety in jeopardy based on this scenario
	10	Learning system is less than ideal and inconvenient for students		18	Calibration must be performed every year
	6	Documentation is poor		11	Legal liability/state laws may determine frequency of calibration
	6	Skill mastery is occurring		7	New policy to save money seems appropriate given the circumstances
	6	Skill mastery is not occurring		7	Maintenance schedule should follow manufacturer suggestions
	6	No learning over time		6	Other places to make cuts, budget should not compromise procedures
	5	Student exposure to equipment meets standard		5	Formal planning and documentation of safety checks is occurring
	5	Supervision is adequate		48	No comment
	4	Isokinetic devices no longer important for athletic trainers			
	3	Learning over time is sufficient			
	3	Supervision of students is lacking			
	46	No comment			
2	20	No written technical standards	7	24	Catalog is a contract, all policies must be published in catalog
	19	No formal review of documents taking place		23	Information is available to public, cannot control printing cycle
	10	Proof of physical examination should be kept with program director		7	Request individual letters be sent informing students of change
	7	General physical examinations no longer adequate		3	Institutional autonomy, final oversight on policy rests with institution
	4	Health questionnaire should be developed		42	No comment
	4	Physical examination policy must equivalent for all students at institution			
	3	Proof should be kept with institution, not program director			
	3	Physical must address psychological health of student	8	11	Documentation of attendance is required
	35	No comment		9	Infection control/exposure policy must be posted at each site/facility
				7	Seminar format is acceptable teaching method for this standard
3	22	Students must be evaluated at middle and end of term		7	Knowledge test of the material covered is required
	20	Evaluations not occurring frequently enough		4	Students must be trained on location at each site/facility
	12	Clinical evaluation must include discussion of didactic performance		4	More training needed, 1 hour session is unsatisfactory
	10	Evaluation must be both summative and formative		3	Safety and prevention must be paramount
	8	Each evaluation must be documented in writing		58	No comment
	3	One meeting over the academic term is sufficient			
	32	No comment	9	36	ATC not physically present to intervene or provide daily supervision
4	22	Course sequencing is irrelevant if students successful on NATABOC		16	Written job description ensures compliance
	22	Course sequencing in this scenario must be improved		8	Program meets established 8:1 Student/ ATC supervision ratio
	15	The institution should have final decision authority, not JRC-AT		4	Sophomore students should have limited assignments
	38	No comment		3	Depends on if student is first aider or athletic training student
				40	No comment
5	49	Survey administration should occur more frequently	10	26	Unacceptable because physician is not an ACI
	10	Survey procedures (cycle of distribution) is appropriate		19	Acceptable as long as final evaluation is performed by ACI
	8	Program should develop quantitative and qualitative instruments		10	Level of physician involvement is a program strength
	30	No comment		42	No comment

Note. Comments/emerging themes appearing less than 3 times not reported. Single respondents often expressed multiple themes. NATABOC, National Athletic Trainers' Association Board of Certification; JRC-AT, Joint Review Committee on Educational Programs in Athletic Training; ATC, certified athletic trainer; ACI, approved clinical instructor.

ardy, which is correct. However, many other respondents failed to recognize the seriousness of the safety issue.

Scenario 7 speaks to the adequacy of the information sources available to prospective students. The information is available to the public, and the program is making a good-faith effort to inform; therefore, the scenario is compliant. Many academic catalogs and bulletins contain a wavier, typically located on the first page, specifying that information contained within the publication is informative in character and thus does not constitute a contract between the student and the institution. However, many respondents believed the academic catalog constituted a contract, which is a myth.

Scenario 8 primarily addresses the issue of annual Occupational Safety and Health Administration (OSHA) training. Further interpretation of this standard reveals that an infection-control policy, including universal precautions, should be available in each clinical setting. As described in the scenario, clinical site policies and procedures for each site were addressed in the hypothetical training session and students were given a handout documenting specific procedures to be followed, ensuring the topic was systematically and uniformly attended to in terms of preventive considerations. Therefore, the scenario is considered compliant, although some respondents called for better forms of documenting attendance at the training session, which is an appropriate recommendation.

Scenario 9 describes availability and settings. Accordingly, the situation described in this scenario is noncompliant due to lack of supervision. In other words, unsupervised student experiences do not count because an approved clinical instructor (ACI) is not physically present to intervene on behalf of the individual being treated regardless of the responsibilities that students are given and their level of preparation. The majority of comments provided for this scenario used language that agreed with the appropriate rationale.

Scenario 10 relates to ACI qualifications. Only ATCs who have completed ACI training are qualified to verify entry-level performance of the proficiencies. Therefore, the medical doctor described in this scenario may continue to teach and evaluate student proficiencies; however, a qualified ACI must assess the final proficiency in stethoscope use. A majority of site visitors responded to this question incorrectly. See Table 4 for a complete listing of all majority themes.

Limitations

It should be noted that actual on-site visits in athletic training are conducted by two people over a period of two or three days and include interviews with administrators, faculty, staff, and students for the purpose of clarifying and confirming previously submitted self-study documents. This study was conducted online using written scenarios to mimic real-life situations; thus, respondents did not

read self-study documents, conduct face-to-face interviews, or work in collaboration with another trained site visitor. In addition, there were only 10 scenarios provided when in fact there are countless situations that a site visitor might encounter during a site visit. Although all members of the site visitor sample had completed training to become a site visitor, several of the respondents either had yet to perform a visit ($n = 12$; 16.7%) or did not respond to the question regarding number of visits conducted ($n = 21$; 22.5%). Therefore, the lack of experience among those conducting actual site visits within the JRC-AT site visitor pool, coupled with the inherent difficulties of an online instrument, may have contributed to the results. In terms of sampling, the respondents may have inadvertently created a biased sample because only those with strong opinions may have participated.

Discussion

Although limitations exist with instruments designed to mimic real-life situations and a single study is insufficient to make major policy changes, the data presented here suggest that current site visitor training in athletic training could be improved for several reasons. First, the overall level of interrater reliability was poor. Put simply, site visitors did not react to identical information in scenarios similarly. Second, the overall average score of the respondents was 68%. On many grading scales, this score is equivalent to a D+. Third, several of the respondent comments contained questionable justification for rating a scenario compliant or noncompliant, and some respondents disagreed with the accepted standards.

Perhaps a knowledge test of the current standards should be given to all site visitors. Continuing education units for site visitations or an appropriate per diem might also provide an impetus for quality improvement by raising the stakes and rewarding good performance. Also, programmed learning systems such as the Examiner Training Home Study Workbooks³² or the NATA Clinical Instructor Educator/Approved Clinical Instructor³³ training program should be considered models for improving consistency among site visitors, although it is unclear whether programmed site visitor training methods would have resulted in greater interrater reliability in the present study.

Alternative methods to improve training might also include interactive CD-ROMs or online modules delivered with streaming video. Moreover, a broader examination of personal bias and the promotion of integrity should be discussed at evaluator training sessions. Several site visitors who participated in the study expressed dissatisfaction with the accreditation process in the comment boxes, in part due to continual procedural changes without adequate training. For instance, correspondence from the JRC-AT^{34,35} indicates that several procedural changes were introduced to site visitors over the past several years such as writing of

reports, yet formal training to accompany the changes did not occur. Likewise, new terminology such as “rejoinder” and changes to the way site visitors receive self-studies were introduced without formal educational in-service or training for site visitors. The confidential Report of On-Site Evaluation form in which site visitors write and submit to the JRC-AT office was changed from a predominately rubric format to a narrative format. Continual procedural changes without training complicate the accreditation process and create frustration. Therefore, efforts to improve public relations should be considered.

Clearly, new methods for ensuring consistent site visitor application of standards should be investigated in light of the potentially devastating consequences that an inaccurate site visit team’s evaluation report may have on an institution, students, faculty, and alumni, especially given that nearly 40% of the subjects sampled did not perform a site visit or did not respond to a question related to the number of site visits conducted. This lack of experience among site visitors in the pool of trained persons may have major implications on an academic program under review. Regrettably, current site visitor training procedures do not address content knowledge or competency. Rather, the JRC-AT training procedure favors an apprenticeship approach whereby novice site visitors learn by working with seasoned visitors. If there is not agreement and accuracy among evaluators, then there will be a great deal of variability across academic programs applying standards for the preparation of athletic trainers.

Conclusions

Findings from this study suggest that improvements in the athletic training accreditation site visitor training process are needed. Accreditation is important to maintaining high standards in higher education and health care; therefore, understanding how site visitors interpret and judge identical information is important for not only being fair but for the continual improvement of health care.

Perhaps this study will inspire others to conduct investigations about the consistency and reliability of accreditation evaluations in other allied health disciplines in the continual effort to improve health care. Recent trends³⁶ among accreditation agencies suggest greater reliance on self-generated annual reports and extended periods of continuing accreditation, which perhaps, over time, could signify decreased dependence on site visits and increased importance of text-based documentation as described here. In other words, as advances in technology become further integrated with outcome assessment techniques, site visitors could eventually be asked to review online self-studies, perhaps like the instrument described in this study. Such potential changes, both real and hypothetical, necessitate the need for well-designed continuing education systems for site visitors and for the people charged with developing and maintaining accreditation standards.

REFERENCES

1. Blauch LE: *Accreditation in Higher Education*. New York, NY: Greenwood Press; 1969.
2. Harclerod FF: *Accreditation: History, Process, and Problems*. Washington, DC: American Association of Higher Education; 1980.
3. Haroian-Guerin G: Letters to the editor: the changing role of accreditation. *Chron Higher Educ: Chron Rev* 2000; Oct 27:B17.
4. Riesman D: Protecting students by voluntary action: regional accrediting associations. *On Higher Education: The Academic Enterprise in an Era of Rising Student Consumerism*. San Francisco, CA: Jossey-Bass; 1980:pp 322–343.
5. Silva MK: *Accreditation, Knowledge, and Strategies of Professionalizing Occupations*. Ann Arbor, MI: UMI Microfilm; 2001.
6. Stark JS, Lattuca LR: *Shaping the College Curriculum: Academic Plans in Action*. Needham Heights, MA: Allyn & Bacon; 1997:pp 18, 232–233.
7. Eaton JS: Regional accreditation reform: who is served? *Change* 2001; 33:38–45.
8. Commission on Accreditation of Allied Health Education Programs: About CAAHEP. Available at: <http://www.caahep.org/>. Accessed 31 Dec 2003.
9. Delforge GD, Behnke RS: The history and evolution of athletic training education in the United States. *J Athl Train* 1999; 34:53–61.
10. Peer KS, Rakich JS: Accreditation and continuous quality improvement in athletic training education. *J Athl Train* 2000; 35:188–193.
11. Proctor RL: *A Survey of Selected Characteristics of Accredited Athletic Training Educational Programs*. Ann Arbor, MI: UMI Microfilm; 2000.
12. McMullan D: Renewing athletic training education: a step into the future. *NATA News* 1996; (Feb):17–18.
13. Koehneke P: Educating the future of the profession. *Athl Ther Today* 2003; (Sep):1.
14. Joint Review Committee on Educational Programs in Athletic Training (JRC-AT): Transition to self-accreditation information. Available at: <http://www.jrc-at.org/trans/index.html>. Accessed 1 Jan 2004.
15. Leverenz LJ: Moving accreditation forward: CAAHEP’s commitment to change. *NATA Educ Council News* 2001; (Jul):9–11.
16. Leatherman C: Failure in national accrediting. *Chron Higher Educ*. [serial online] 1995; Jun 9. Available at: <http://chronicle.com/che-data/articles.dir/articles-41.dir/issue-39.dir/39a01901.htm>. Accessed 30 June 2001.
17. Dill DW, Massy WF, Williams PR, et al: Accreditation and academic quality assurance: can we get there from here? *Change* 1996; 28: 17–24.
18. Greenberg M: America’s colleges should rank themselves. *Chron Higher Educ: Opin Arts* 2000; Jun 16:B10.
19. Honan WR: Some say college accreditation is out of control. *New York Times* [serial online]. 1998; Nov 11. Available at: <http://www.nytimes.com/library/1198educaccreditation.html>. Accessed 30 June 2001.
20. McMurtrie B: Assessing the group that assesses accreditation. *Chron Higher Educ: Money Manage* 1999; Nov 12:A41.
21. Foster DT, Cuppett M, Gardner G, et al: *Manual for the Interpretation of Standards and Guidelines for an Accredited Educational Program for the Athletic Trainer*. Centennial, CO: Joint Review Committee on Educational Programs in Athletic Training; 2001.
22. Joint Review Committee on Educational Programs in Athletic Training: *JRC-AT Site Visitor Workshop Agenda*. Centennial, CO: Joint Review Committee on Educational Programs in Athletic Training; 2001.
23. Joint Review Committee on Educational Programs in Athletic Training (JRC-AT): An investigation of inter-rater reliability among athletic training site visitors [online survey Web site]. Available at: <http://muweb.millersville.edu/~jwimer/onsite.htm>. Accessed 11 Jan 2002.

24. Lincoln YS, Guba EG: *Naturalistic Inquiry*. Newbury Park, CA: Sage Publications; 1985.
25. Cunningham G: *Educational and Psychological Measurement*. New York, NY: Macmillan Publishing Co; 1986.
26. Morrow J, Jackson A, Disch J, et al: *Measurement and Evaluation in Human Performance*. Evanston, IL: SYSTAT, Inc; 1995.
27. Shrout PE, Fleiss JL: Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; 86:420–428.
28. Heiman GW: *Basic Statistics for the Behavioral Sciences*. Boston, MA: Houghton Mifflin Co; 2000.
29. Statistical Package for the Social Sciences Version 10.0. Chicago, IL: SPSS, Inc; 2000.
30. Bilgrave DP, Deluty RH: Religious beliefs and political ideologies as predictors of psychotherapeutic orientations of clinical and counseling psychologists. *Psychother Theory Res Pract Train* 2002; 39:245-60.
31. Appelbaum LD: The influence of perceived deservingness on policy decisions regarding aid to the poor. *Political Psychol* 2001; 22: 419–442.
32. National Athletic Trainers' Association Board of Certification, Inc: *NATABOC Examiner Training Home Study Workbook*. Omaha, NE: National Athletic Trainers' Association Board of Certification, Inc; 1999.
33. National Athletic Trainers' Association: *Clinical Instructor Educator Seminar*. Dallas, TX: National Athletic Trainers' Association; 2000.
34. Joint Review Committee on Educational Programs in Athletic Training: Conflict of Interest Form for Potential Fall 2002 and Spring 2003 site visits. Centennial, CO: Joint Review Committee on Educational Programs in Athletic Training; April 29, 2002.
35. Joint Review Committee on Educational Programs in Athletic Training: Letter from the Chair: Peter Koehneke, Spring 2002. Centennial, CO: Joint Review Committee on Educational Programs in Athletic Training; April 29, 2002.
36. Commission on Accreditation of Allied Health Education Programs: Memorandum from Kathleen Megivern, CAAHEP Executive Director. Chicago, IL: Commission on Accreditation of Allied Health Education Programs; May 15, 2002.

APPENDIX. STANDARDS AND SCENARIOS

Listed are the 10 standards and the 10 scenarios from the online instrument.

1. Standard: IB3b(1) Physical Resources—Therapeutic Modalities and Rehabilitation

Therapeutic modalities and rehabilitation equipment shall be available for instructional purposes. Such equipment shall include items identified in the “NATA Athletic Training Educational Competencies.” Therapeutic modalities and rehabilitation equipment appropriate to the clinical setting shall be available for clinical education purposes.

Scenario 1

Countryside State College, an institution of 1,200 students, is seeking initial accreditation of its Athletic Training Education Program (ATEP). Students learn about isokinetic equipment procedures in a series of classroom lectures. After the lectures, students are exposed to isokinetic equipment in a one-hour lab at a physical therapy clinic 15 minutes from campus. While there, an Approved Clinical Instructor (ACI) evaluates and documents student mastery of isokinetic procedures. The ACI then instructs the students that they should visit the facility at their convenience over the next several semesters to practice using isokinetic equipment and to prepare for a follow-up evaluation before graduation. Students must sign an attendance sheet, which is forwarded to the Program Director (PD) indicating that they practiced using the equipment over these semesters. Based on the information provided above, rate this program in accordance to the Standard.

2. Standard: IC3 Health

Procedures shall be established and implemented to determine that the students' physical and mental health will

permit them to meet the established written technical standards of the program.

Scenario 2

Located in Capital City, Tower State University is the state's flagship institution with a student enrollment of nearly 30,000. Experiencing a long and successful history of NATA approval/CAAHEP accreditation, the ATEP enrolls approximately 80 freshmen-senior level students. Admission to the University requires that all students submit proof of a general physical examination and appropriate immunizations (i.e., measles, mumps, rubella, tetanus, and diphtheria). In addition to the University requirement, the PD requires that her Athletic Training students provide documentation (or formal declination) of Hepatitis B vaccinations. The HBV documentation is the only medical record kept on file in the PD's office. Based on the information above, rate this program in accordance to the Standard.

3. Standard: IC2a(1) Evaluation of Students

Evaluation methods/systems shall be implemented for both didactic and supervised clinical education components. They shall be employed frequently enough to provide students and program officials with timely indications of the students' progress and academic standing and to serve as a reliable indicator of the effectiveness of course design and instruction.

Scenario 3

Athletic training students at Oceanside University are required to meet with their Clinical Instructors (CI) at the completion of each clinical course for the purpose of reviewing their overall performance. Over the years, this meeting has developed into an opportunity to discuss the overall

quality of the student's clinical performances as well as provide them with specific recommendations for improvement based on course objectives. Based on the information above, rate this program in accordance to the Standard.

4. Standard: IIA2b Instructional Plan

Instruction shall follow a plan that documents: a logical progression of didactic study and clinical experience opportunities.

Scenario 4

During your site visit to Tarmac University, you and the team chair observe that Therapeutic Exercise is offered to junior-level students in the same semester as the Kinesiology course. The team chair suggests that the course sequencing is academically unsound because students receive formal instruction in rehabilitation techniques before learning kinesiological concepts. However, program graduates that you interviewed appear to be successful, passing the NATABOC exam at a rate higher than the national average. Moreover, employer follow-up data about the alumni, especially those in the non-traditional setting, are exceptional. Based on the information above, rate this program in accordance to the Standard.

5. Standard: IE1 Outcomes

Programs shall routinely secure sufficient qualitative and quantitative information regarding the program graduates to demonstrate an ongoing evaluation of outcomes related to the educational and program objectives.

Scenario 5

Since its last site visit in 1996, Greenlawn State College has prepared and distributed one "alumni survey" and one "employer survey." The intent of both surveys (sent and analyzed in 1998) was to determine student satisfaction and to solicit recommendations for the ATEP. During your review, you find both surveys to be well designed, as well as several instances whereby the ATEP used data from the surveys for program improvement. Based on the information provided above, rate this program in accordance to the Standard.

6. Standard: 1D1g Fair Practices

The health and safety of patients, students and faculty associated with the educational activities of the students shall be adequately safeguarded.

Scenario 6

Located in a remote part of the state, General College's primary mission is to serve students from low socioeconomic backgrounds. For many years, tuition at General College was

among the lowest. However, in response to recent economic downturns and low enrollments, state legislators have mandated budgetary cuts and a hiring freeze. In order to save \$725 per year, the ATEP decided to postpone a service contract in which formal modality calibration and safety checks were conducted by an off-campus company. As a result, formal checks will occur every other year, and informal checks as needed. Based on the information provided above, rate this program in accordance to the Standard.

7. Standard: 1C1b Admission Policies and Procedures

Program admission criteria shall be clearly defined and published in the official institutional academic documents and other public media.

Scenario 7

At the request of the new Academic Dean, program officials at Lakefront University decided to change the grade point average used for formal admission into the ATEP from 2.5 to 2.75. Regrettably, this approved change will not be published in the official Academic Catalog for another two years due to the University's printing cycle. The PD, however, has agreed to post the new change on the ATEP's website, in the brochure, and in recruitment letters. Based on the information provided above, rate this program in accordance to the Standard.

8. Standard: 1D1h Fair Practices

The program shall comply with Occupational Safety and Health Administration blood-borne pathogen requirements. Education in pathogen and infection control shall be provided annually.

Scenario 8

Each year, as part of an annual in-service conducted by the campus health clinic, students in the ATEP at Academy College complete a one-hour educational program on prevention of blood-borne pathogen transmission and universal precautions, including protective equipment. Near the end of the in-service, students are given a handout that addresses emergency policies and procedures in case a student accidentally suffers exposure in one of the seven off-campus clinical settings used by the ATEP. Based on the information provided above, rate this program in accordance to the Standard.

9. Standard: IIA1g Description of the Program

Supervised clinical experiences shall involve the daily personal contact and supervision between the clinical instructor and the student in the same clinical setting. The instructor shall be physically present in order to intervene on behalf of the individual being treated.

Scenario 9

Fairview College, which participates in NCAA Division III athletics, employs 3 full-time ATCs, each of whom also serve as ACIs for the ATEP. During your site visit, you determine that sophomore- and junior-level athletic training students are placed at the baseball, softball, and soccer fields while two ATCs cover football and one ATC is in the athletic training room readily available should the students at baseball, softball or soccer need him. The PD points out that each of the ATCs is in walkie-talkie contact of students, and that each of the ATCs could respond to the respective fields within 2–4 minutes via the golf cart. Furthermore, the PD states that each student must pass Fairview's advanced first aid and CPR course before being assigned to clinicals, and that written job descriptions outline the roles and responsibilities of students. Based on the information provided above, rate this program in accordance to the Standard.

10. Standard: 1B1c(1)(a) Other Instructional Staff ACI Responsibilities

An ACI shall be a faculty, staff, or adjunct allied health or medical community member of the sponsoring institution

or affiliates who provides instruction and/or evaluation of students in the clinical proficiencies of the athletic training educational program. An ACI shall perform psychomotor and/or clinical proficiency instruction and evaluation at some point during the educational experience. Evaluation of the proficiency shall be done in a one-on-one basis as determined by the institution. The ACI shall also be involved in the learning over time continuum during the clinical experience.

Scenario 10

Jane Smith, MD, FACSM, has served as the team physician at West Coast University for the past 10 years. An active participant in the ATEP, Dr. Smith shares her knowledge through ongoing informal discussions in the athletic training room and by serving as a part-time classroom instructor in a course entitled General Medical Conditions. As part of the clinical instruction plan, Dr. Smith has agreed to teach and evaluate student proficiencies in stethoscope usage (i.e., normal breath sounds, normal heart sounds, and normal bowel sounds). Based on the information provided above, rate this program in accordance to the Standard.