A Risk Stratification Tool to Assess Commercial Influences on Continuing Medical Education

BARBARA E. BARNES, MD, MS; JEANNE G. COLE, MS; CATHERINE THOMAS KING, BBA; REBECCA ZUKOWSKI, RN, MSN; TRACY ALLGIER-BAKER, BS; DORIS McGARTLAND RUBIO, PHD; LUANNE E. THORNDYKE, MD

Introduction: Heightened concerns about industry influence on continuing medical education (CME) have prompted tighter controls on the management of commercial funding and conflict of interest. As a result, CME providers must closely monitor their activities and intervene if bias or noncompliance with accreditation standards is likely. Potential for industry influence can be difficult to assess at a stage of the planning process when mitigation strategies can assure balance and content validity. Few tools exist to aid providers in this regard.

Methods: A 12-item instrument was designed to assess risk for commercial influence on CME. To determine reliability and validity, a cohort of experienced CME professionals applied the tool to standardized “cases” representing CME activities in the early stages of planning. Results were compared with the experts’ assignment of the same cases to one of four risk categories. A survey of study participants was conducted to ascertain usefulness and potential applications of the tool.

Results: Analysis demonstrated strong intraclass correlation across cases (0.90), interrater reliability (94%), and correlation between assessment of risk with and without the tool (Spearman coefficient, 0.93, p < 0.01; weighted kappa, 0.59). Participants found the tool easy to use and of potential benefit to their CME office.

Discussion: The Consortium for Academic Continuing Medical Education (CACME) risk stratification tool can help CME providers identify activities that must be closely monitored for potential industry influence, remain aware of factors that place programming at risk for noncompliance with accreditation standards, and substantiate the allocation of resources by the CME office.

Key Words: continuing medical education, commercial interests, standards for commercial support, risk stratification tool, risk assessment, ethical standards, Consortium for Academic Continuing Medical Education

Introduction

Increasing levels of commercial funding for continuing medical education (CME) in the United States have raised concerns about the effect of industry on the quality and scientific balance of physician education as well as the independence of the CME enterprise. In 2006, 61% of CME revenue in the United States was derived from commercial sources, such as the pharmaceutical and medical device industry. Given that CME participants often have difficulty determining the difference between commercial bias and expert personal opinion, educators must play an active role in assuring that programming is balanced and evidence-based. Standards promulgated by various accrediting organizations, including the Accreditation Council for Continuing Medical Education (ACCME), the American Academy of Family Physicians, and the American Medical Association, hold CME providers accountable for the scientific integrity of their programs. Despite the expenditure of considerable administrative resources to adhere to regulatory requirements, it is discouraging to recognize that 20% of ACCME-accredited organizations are in noncompliance with one or more elements of the Standards for Commercial Support (SCS). CME providers must develop rigorous methods to recognize, manage, and assess commercial influence. Tools developed for this purpose, such as surveys of participants’ perceptions of bias, are customarily used at the end of an activity when it is too late for providers to intervene.
Risk stratification is a statistical process employed in medicine to determine factors associated with adverse clinical outcomes, so that practitioners can develop targeted interventions to mitigate their impact.11,12 For example, the identification of risk factors for coronary artery disease (e.g., hypertension, diabetes, hyperlipidemia, cigarette smoking) is used as a basis for initiating a variety of lifestyle modifications and medical interventions that can reduce morbidity and mortality rates.11 This approach has been employed in numerous other clinical conditions.13–16 Application of these principles to the CME enterprise offers the opportunity to prospectively identify factors that may adversely affect the balance and educational integrity of programming. Interventions can then be developed to maintain compliance with regulatory standards and, more importantly, provide effective learning experiences.

In 1998, the Consortium for Academic Continuing Medical Education (CACME)2 began developing a risk stratification tool to assess potential for commercial influence. Using nominal group and Delphi techniques, four major factors contributing to risk were identified: (1) the nature of the activity’s content, (2) the amount and characteristics of commercial support, (3) the level of control of the CME office in planning and financial management, and (4) the potential influence of commercial entities on various stakeholders, such as course directors. Criteria for assessment of each of these factors were refined through pilot testing, resulting in a 12-item instrument (Appendix). Numerical weighting of these items permitted calculation of an overall risk score, and ranges of scores were grouped into categories of low, moderate, high, and very high risk, with management and oversight processes developed for each level. Over time, results obtained through use of the tool were compared with actual monitoring and compliance issues of CACME activities, resulting in ongoing refinement of item weighting and operational definitions. Participating schools have routinely used the risk tool at the beginning of the planning process for all activities except regularly scheduled conferences. Strategies to mitigate the risk identified by the tool have included denial of certification, enhanced monitoring, more rigorous resolution of conflict of interest, and increased involvement of the CME office in educational design and logistical support. Ongoing quality assurance reviews have demonstrated consistent compliance of activities with the ACCME’s Standards for Commercial Support.

Methods

The usefulness of the tool in the CACME institutions stimulated interest in its dissemination to other CME providers. In order to enhance understanding of the value of the instrument, a study was conducted to determine whether a tool based on the principles of risk stratification, applied early in the planning process, could reliably identify issues that place CME activities at risk for commercial influence. To address this question and gain some understanding of the validity of the tool, we asked experienced CME providers, all of whom were knowledgeable about ACCME standards, to assign a set of CME activity “cases” to one of four risk categories (low, moderate, high, and very high). We compared these assessments to scores obtained when the same individuals applied the risk tool to the cases (FIGURE 1).

Development of Standardized Cases

To provide a standardized evaluation environment, the “cases” were developed to represent data commonly available early in the CME planning process. Although these cases were based on actual CACME programs, some characteristics were modified to create representation among the four risk categories (low, moderate, high, and very high). Cases were placed into a structured format that included a general description of the proposed activity and details about the budget, funding sources, management, meals, social events, and relationships with industry (information that would normally be available early in the planning process). Every option for each of the 12 items included in the risk stratification tool was represented in at least two of the cases included in the study. Preliminary evaluation of the cases for test-retest reliability was conducted using CACME deans, administrators, and staff who rated the cases on three separate occasions using the risk stratification tool. As a result of this exercise, modifications were made to improve the clarity and completeness of the information provided, yielding a total of 36 cases (9 in each risk category).

Design

On the basis of power calculation, it was determined that at least 12 ratings per case using the tool would be required to
achieve statistically meaningful results. Study participants, who were each assigned 24 cases representing various levels of risk for commercial influence, were divided into two groups. Group 1 participants were sent the cases along with the risk stratification tool (without the associated numerical ratings), operational definitions for each of the 12 items, and two “practice cases” to familiarize them with use of the instrument. The individuals were asked to apply the tool to the cases and return all materials to the study team. Three weeks later, participants received the same cases and were asked to assign them to one of four risk categories. Operational definitions of each category were provided, along with two “practice cases.” Group 2 received the cases along with the risk categories first and then repeated the exercise 3 weeks later using the risk stratification tool. After the second assessment, all participants received a 17-item survey using a 5-point Likert scale to assess perceptions of benefits, usability, and limitations of the tool.

Study Participants
A convenience sample of 22 CME professionals from across the United States was invited to participate. Of these, 18 agreed. Participants had an average of 13.8 years of experience in CME and worked in a variety of institutional settings. Eleven were ACCME surveyors and 2 were ACCME staff. None of them had ever seen or used the risk stratification tool. A $50 honorarium was offered to each participant for completion of the study.

Statistical Analysis
Descriptive statistics were used to report the characteristics and tabulate the responses of study participants. Assessment of the tool’s reliability was determined by calculation of the intraclass correlation coefficient across the 36 cases. Interrater reliability was also calculated and expressed in terms of a mean value for each item and for all items.

To assess the tool’s validity, total risk stratification scores were correlated with the assigned risk categories across the 36 cases. Results were expressed in terms of the Spearman correlation coefficient, which is a correlation between the level of risk as determined by the tool and the assigned risk categories across all cases. Weighted kappa, which is ordinarily considered to be a measure of consistency assessing agreement among raters extending beyond chance, was employed to assess validity. Validity was determined by consistency between the total risk score and the participants’ assignments to a risk category without benefit of the tool.

To assess the tool’s usefulness, Likert scores were tabulated for each survey item and the mean score was calculated for each. Analyses were performed with Stata (StataCorp, College Station, TX) and SAS (SAS Institute, Cary, NC) in the Data Center at the University of Pittsburgh Center for Research on Health Care.

Results
All participants completed the study. One set of data was incomplete and therefore unusable. The reliability of the tool among the different raters was strong. The intraclass correlation coefficient was 0.90, based on 36 cases (TABLE 1). Interrater reliability was calculated by the average agreement for each item. Results indicated strong intrarater reliability of the tool, with an average agreement of 94% for all items (standard deviation = 9.27). The two items with the lowest degrees of reliability were the delegation of logistical responsibilities to another party (89%) and the level of involvement of a commercial supporter in suggesting topics (90%). Each of the remaining items showed greater than 90% agreement.

In order to assess association and bias, we calculated the percentage agreement across raters for each of the cases. Across all 36 cases, the percentage agreement ranged from 85% to 99%, with an average agreement of 94%. This result demonstrates that the differences in scores were not attributable to the differences among the raters.

The validity of the tool was assessed using two methods. First, we calculated the correlation across the 36 cases between the tool and the subjective rating of the cases. Because of the small sample size, the four risk categories were collapsed to two, combining the low and moderate levels (low risk) and the high and very high levels (high risk). Using this methodology, the Spearman correlation coefficient was 0.93 ($p < 0.01$), indicating a good correlation between the categories determined through use of the risk score and those assigned by the raters. Second, we estimated a weighted kappa to determine the extent to which

<table>
<thead>
<tr>
<th>TABLE 1. Summary of Statistical Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Reliability</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Average agreement</td>
</tr>
<tr>
<td>Intraclass correlation coefficient, single rater</td>
</tr>
<tr>
<td>94% (standard deviation = 9.27)</td>
</tr>
<tr>
<td>$p = 0.000$</td>
</tr>
<tr>
<td>$(n = 36)$</td>
</tr>
<tr>
<td>Spearman correlation coefficient</td>
</tr>
<tr>
<td>0.932</td>
</tr>
<tr>
<td>$(n = 36)$</td>
</tr>
<tr>
<td>Validity</td>
</tr>
<tr>
<td>Weighted kappa</td>
</tr>
<tr>
<td>0.5866</td>
</tr>
<tr>
<td>$(n = 408)$</td>
</tr>
</tbody>
</table>

DOI: 10.1002/chp
the overall score obtained through use of the tool is consistent with the participants' categorical assignments. To estimate the kappa, 24 ratings for each case were used (12 ratings using the tool and 12 from category assignment), yielding a sample size of 408. The correlation between the ratings using the tool and 12 from category assignment indicated that the tool has clear operating definitions (mean score of 4.0 on a 5-point Likert scale), can be helpful in allocating resources to high-risk activities (4.20), has potential use for teaching staff about standards for commercial support (4.47), and can assist CME professionals in thinking about issues that affect compliance with the Standards for Commercial Support (4.67). Fourteen of 15 respondents agreed or were neutral about the potential use of the tool by the ACCME to select files for accreditation surveys, but 5 indicated that they would not want to share such information with the ACCME.

**Discussion**

Under increased pressure to prevent industry influence, CME providers must develop mechanisms for consistently identifying and managing vulnerable activities. Risk stratification methods have been employed in multiple areas of health care and in other disciplines. However, a review of the literature reveals only one article in which these methods were used to assess educational outcomes. This study demonstrates a way to apply risk stratification principles to assess the potential for a CME activity to be influenced by industry, and therefore noncompliance with ACCME Standards for Commercial Support. Our results indicate that the CACME risk stratification tool provides a mechanism to estimate risk that is reliable, potentially valid (as judged by the standard of expert assessment of risk), and useful to CME providers.

Two principles of risk assessment are particularly relevant to CME. One is the prediction of risk, with an emphasis on identification of risk factors and stratification of overall risk. The other is the management of risk, with an

<table>
<thead>
<tr>
<th>TABLE 2. Participant Survey Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The risk stratification tool makes me think about issues that affect compliance with the Standards for Commercial Support.</td>
</tr>
<tr>
<td>2. The risk tool is difficult to use.</td>
</tr>
<tr>
<td>3. The operational definitions for the elements of the risk stratification tool are clear and easy to understand.</td>
</tr>
<tr>
<td>4. Use of the risk stratification tool is time consuming.</td>
</tr>
<tr>
<td>5. Routine use of the risk stratification tool would help my CME department decide which activities to certify.</td>
</tr>
<tr>
<td>6. Routine use of the risk stratification tool would assist my CME office in allocating resources to activities at higher risk of noncompliance with ACCME requirements.</td>
</tr>
<tr>
<td>7. The risk stratification tool would be useful for teaching staff about issues associated with the Standards for Commercial Support.</td>
</tr>
<tr>
<td>8. It would take a long time to train my staff to use the risk stratification tool.</td>
</tr>
<tr>
<td>9. Results obtained through the use of the risk stratification tool would be useful to the ACCME in deciding which files to review during accreditation surveys.</td>
</tr>
<tr>
<td>10. I would not want to share the risk information obtained through use of the risk stratification tool with the ACCME.</td>
</tr>
<tr>
<td>11. Information obtained through use of the risk stratification tool would be helpful in justifying to course directors and joint sponsors why certain requirements of the CME office are necessary.</td>
</tr>
<tr>
<td>12. Use of the risk stratification tool would help to justify why the CME department assesses higher fees for certain activities.</td>
</tr>
<tr>
<td>13. The concept of a risk stratification tool is potentially very valuable but the instrument provided is not adequate.</td>
</tr>
<tr>
<td>14. Certain key elements are missing from the risk stratification tool that was provided.</td>
</tr>
<tr>
<td>15. Use of the risk stratification tool adds little to my own subjective assessment of an activity’s potential risk for noncompliance with the ACCME Standards for Commercial Support.</td>
</tr>
<tr>
<td>16. I would like to use the risk stratification tool routinely in my CME office</td>
</tr>
<tr>
<td>17. Information obtained through use of the risk stratification tool would be helpful in explaining to industry representatives why certain requirements of the CME office are necessary.</td>
</tr>
</tbody>
</table>

*Note: 1 = Strongly disagree; 2 = Disagree; 3 = Neutral; 4 = Agree; 5 = Strongly agree.*
emphasis on risk reduction.\textsuperscript{17} The CACME tool addresses both principles. First, by helping CME professionals identify factors that can contribute to noncompliance with standards, the tool prompts data collection about specific program issues that are critical to maintenance of compliance. Second, once risk factors have been identified, planners can develop strategies to mitigate these factors. The higher the category of risk, the greater the need for “risk management,” including such actions as stringent monitoring, review of documentation, and allocation of resources and personnel to ensure compliance. Thus, the CACME risk stratification tool provides a prediction that can influence the decision about whether to certify a given CME activity and can guide management throughout planning and development. In addition, the tool may be useful for training CME staff and educating activity directors, joint sponsors, funders, and other individuals who contribute to the implementation process. Finally, compilation of aggregate risk scores and data about individual activities allows providers to evaluate their overall CME program.

CACME’s work demonstrates the value of multiinstitutional collaboration in addressing complex issues associated with educational planning and certification. In developing the tool, the iterative process we employed drew on the expertise of leaders from all member schools. It is unlikely that any single individual or institution could have developed such a tool in isolation. Applying the tool across the consortium allowed us to monitor the performance of the individual members and the consortium as a whole. This, in turn, expanded CACME’s ability to use the tool for decision making, program planning, and evaluation.

Limitations

This study used CME experts to determine a “gold standard” of potential risk (assignment of cases to a risk category) against which the tool was measured. The ideal design to test our instrument would be a naturalistic study in which activities were assessed with the tool and then allowed to proceed without any intervention by the accredited provider (analogous to the natural course of disease), with a follow-up evaluation to determine whether commercial influence or noncompliance with accreditation standards was present. However, this situation does not occur in the real world because accredited providers do intervene to assure compliance. Determination of risk factors in the clinical setting is customarily accomplished through the use of large epidemiologic studies. Such an approach is beyond the scope of this study. We relied on participants’ experience as providers and their knowledge of accreditation requirements to make accurate determinations of potential risk and recognize that this approach limits conclusions about the validity of the tool. However, the validity of the instrument demonstrated in this study has been supported by experience within CACME that indicates frequent compliance issues in activities with high risk scores.

Another concern is that simulated cases were used rather than real activity files. Although the cases were based on actual CME programs, several were modified to represent high-risk activities and to ensure a balanced number of activities in all four risk categories. While this method may not reflect the actual case mix of an accredited CME provider, we wanted to emphasize detection of “outlier” high-risk activities; although they are uncommon, it is critically important to identify them in order to maintain compliance. We also recognize that the information was provided in a structured and consistent fashion, addressing the elements of the activity appropriate to the items contained within the tool. In the day-to-day practice of CME, some data may not be readily available early in the planning process and the individuals using the tool may not be knowledgeable about some of the factors (such as previous experience with a joint sponsor or logistical partner). As a result, the reliability of the tool may be lower than observed in this study.

The size and characteristics of our study sample are relatively limited by the number of available CME experts. Because of the size, the four risk categories were collapsed into two for purposes of one of the measures of validity, limiting conclusions about the discriminatory capacity of the tool across the wide range of risk. The use of seasoned CME professionals in this study also raises questions about the ability to generalize use of the tool among CME providers with varying backgrounds and experience. Long-standing implementation within the four CACME institutions indicates that staff at all levels can use the instrument with reliable results. In addition, study participants indicated that the tool would be useful in their offices.

It is possible that some important risk factors are not included in our tool. During development, some factors that were difficult to quantify or operationally define were excluded. Additional use and evaluation of the tool may permit incorporation of such factors.

Conclusions

The CACME Risk Stratification Tool supports the prospective identification of CME activities that may be at risk for commercial influence. The 12-item tool is easy to use and allows users to categorize each CME activity’s potential for noncompliance as low, medium, high, or very high. The results of a case-based crossover study confirm that the tool is reliable and may have predictive value in terms of risk for commercial influence. A survey of users indicates that it would be useful for objectively assessing a CME program in the planning stage. In addition, the tool can assist in teaching staff about standards for commercial support of CME activities, help individuals identify issues that affect compliance with the standards, and support administrators in allocating resources to mitigate the risks of noncompliance. Through the use of this tool, CME providers can define rational and consistent strategies for certification and management of activities. Such an approach may be
Lessons for Practice

- It is incumbent upon CME providers to implement mechanisms to assure that activities are free of commercial influence and bias.
- Risk stratification provides a mechanism to identify commercial influence in CME and mitigate risk for noncompliance with ACCME Standards for Commercial Support.
- A risk stratification tool has proved reliable, valid, and useful in identifying CME activities at risk for commercial influence.

particularly beneficial in large, decentralized programs that delegate many of these responsibilities to individuals outside the CME office.

There is increasing public demand for transparency and management of relationships between health care organizations and industry.19,20 It is critically important for CME providers to assure that educational programming is evidence-based, balanced, free of commercial influence, and effective. Instruments such as the risk stratification tool provide an objective and systematic means to assess the level of involvement by external entities prospectively in order to allocate appropriate resources to foster compliance with regulatory standards and preserve educational integrity.

Acknowledgments

This project was submitted to the University of Pittsburgh Institutional Review Board and was granted exempt status. Funding for this project was provided by the Consortium for Academic Continuing Medical Education (CACME) and through the Alliance for CME’s Collaboration Award received by CACME in 2001. There was no industry funding for the project. We thank Dr. Timothy Brigham, Dr. Dennis Lott, Ms. Sharon Maddox, Dr. Geno Merli, Dr. Arnold Meyer, Dr. Robert Smedley, and Mr. Derek Warnick for their collaboration on this project, as well as our colleagues who participated in the study.

Notes

a. CACME members consisted of Jefferson Medical College, Penn State College of Medicine, Temple University, and the University of Pittsburgh. In 2005, CACME was disbanded as an accredited consortium, but the four schools continue to collaborate on this and other projects.

b. Traditionally kappa is considered a measure of consistency. However, when one of those measures is the standard by which the other is judged (in this case, the expert judgment of risk), one can infer the validity of the second measure to the extent that it is consistent with that standard.

References

Appendix: CACME Risk Stratification Tool

NAME OF ACTIVITY: _______________________________________________________

1. Joint sponsorship
   Activity is directly sponsored, no joint sponsors: 0
   Activity is jointly sponsored and all joint sponsors are nonprofit organizations: 1
   Activity is jointly sponsored and some or all joint sponsors are for-profit organizations: 2

2. Experience with jointly sponsoring organizations
   Activity is directly sponsored, no joint sponsors: 0
   Positive experience with all jointly sponsoring organization(s): 1
   No prior experience with one or more of the jointly sponsoring organizations: 1
   Some negative experiences with one or more of the jointly sponsoring organizations: 3

3. Commercially supported activity with a standard curriculum delivered at multiple locations
   No commercial support: 0
   Commercial support but the activity does not feature multiple presentations of the same curriculum at different locations: 0
   Commercially supported activity with multiple presentations of the same curriculum at different locations: 3

4. Responsibility for course logistics
   Handled entirely by staff from the CME office: 0
   Some or all responsibilities delegated by the CME office to one or more entities: 1
   Some or all responsibilities delegated by a joint sponsor to one or more entities: 2
   Some or all of logistics performed by an organization suggested or chosen by a commercial supporter: 3

5. Experience with the entity(ies) external to the CME office responsible for some or all of logistics
   Not applicable (all logistics handled by the CME office): 0
   Positive experience working with the entity(ies): 1
   No experience working with the entity(ies): 1
   Negative experience with the entity(ies): 3

6. Responsibility for funds management
   The CME activity has no income or expenses: 0
   Funds management handled entirely by staff from the CME office: 0
   Some or all funds management handled by a not-for-profit entity outside the CME office: 2
   Some or all funds management handled by a for-profit entity external to the CME office: 3

7. Number of commercial supporters
   No commercial support for the course: 0
   Two or more commercial supporters: 2
     One commercial supporter with whom the accredited provider has worked and has had good experiences: 3
     One commercial supporter with whom the accredited provider has never worked: 4
     One commercial supporter with whom the accredited provider has worked and has had negative experiences: 5

8. Anticipated amount of commercial support as a percentage of the anticipated total revenue
   No commercial support: 0
   Up to 50% of course revenue will be from commercial support: 1
   51% to 99% of course revenue will be from commercial support: 3
   100% of course revenue will be from commercial support: 4

9. Level of involvement of any commercial supporter(s) (choose all that apply): MAXIMUM SCORE OF 3
   No commercial support: 0
   Commercial support but no involvement of any commercial supporters with the logistical or educational aspects of the activity: 1
     Assistance from the commercial supporter with marketing: 1
     Suggestion from the commercial supporter for course location: 1
     Recommendation from the commercial supporter of potential participants: 1
     Recommendation from the commercial supporter of the topic area: 1
     Recommendation from the commercial supporter of one or more speakers: 1

10. Anticipated amount of exhibit funds as a percentage of anticipated total revenue of the activity
    There will be no exhibit revenue: 0
     Up to 50% of course revenue will be from exhibit revenue: 1
     51% to 100% of course revenue will be from exhibit revenue: 2

11. Discussion of experimental and/or off-label uses
    Primary intent does not involve the discussion of experimental and/or off-label uses: 0
    Primary intent does involve the discussion of experimental and/or off-label uses: 2

12. Relationships between the course director(s) and industry that might affect the scientific balance of the content
    No relationships exist between course director(s) and industry that are likely to affect the scientific balance: 0
    Relationships exist between course director(s) and industry that are likely to affect the scientific balance of the activity: 2

TOTAL SCORE _______________________________________________________

RISK CATEGORY:
   __ LOW (≤ 2)
   __ MODERATE (3 TO 11)
   __ HIGH (12 TO 19)
   __ VERY HIGH: (≥ 20)